

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 2月17日
Date of Application:

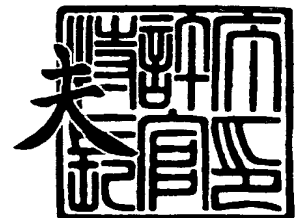
出願番号 特願2003-038201
Application Number:
[ST. 10/C]: [JP 2003-038201]

出願人 日本電気株式会社
Applicant(s):

2003年12月11日

特許庁長官
Commissioner,
Japan Patent Office

今井 康



出証番号 出証特2003-3102691

【書類名】 特許願

【整理番号】 35001174

【提出日】 平成15年 2月17日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/173

【発明者】

【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

【氏名】 細見 岳生

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100088328

【弁理士】

【氏名又は名称】 金田 暢之

【電話番号】 03-3585-1882

【選任した代理人】

【識別番号】 100106297

【弁理士】

【氏名又は名称】 伊藤 克博

【選任した代理人】

【識別番号】 100106138

【弁理士】

【氏名又は名称】 石橋 政幸

【手数料の表示】

【予納台帳番号】 089681

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9710078

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 マルチプロセッサシステム及びデバイス共有方法

【特許請求の範囲】

【請求項 1】 少なくとも 1 つの CPU を備えた複数の CPU セルを有し、前記 CPU セルが複数のグループに分割され、該グループ毎に異なるオペレーティングシステムで動作するマルチプロセッサシステムであって、

前記複数の CPU セル間で共用されるデバイス、

並びに前記デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備え、前記 CPU セルからコマンドを受信すると、該コマンドの発行元に対応するデバイス管理情報を検索し、前記コマンドにより更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させるデバイス管理部を備えた、前記 CPU セルとネットワークを介して接続されたデバイスセルを有するマルチプロセッサシステム。

【請求項 2】 前記デバイス管理情報は、

前記複数の CPU セルにそれぞれ対応して設けられた請求項 1 記載のマルチプロセッサシステム。

【請求項 3】 前記デバイス管理情報は、

前記グループの数、及び前記 CPU セルの数に一致しない、任意の数だけ設けられた請求項 1 記載のマルチプロセッサシステム。

【請求項 4】 前記 CPU セルは、利用可能な前記デバイスの情報をテーブル形式で保持し、

システム立ち上げ時に自機に割り当てられたデバイスを優先的に利用する請求項 1 記載のマルチプロセッサシステム。

【請求項 5】 前記デバイスセルは、

複数の同じデバイスを備え、

前記デバイス管理部は、

前記デバイス管理情報で指定された処理を、前記複数のデバイスのうちの任意のデバイスに実行させる請求項 1 乃至 4 のいずれか 1 項記載のマルチプロセッサシステム。

【請求項 6】 前記 CPU セルは、

前記 CPU から発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、

前記デバイスセルは、

前記コマンドを分解して複数の命令を抽出するコマンド解析部を備え、

前記デバイス管理部は、抽出された複数の命令により更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させる請求項 1 乃至 5 のいずれか 1 項記載のマルチプロセッサシステム。

【請求項 7】 前記 CPU セルは、自機が所属する前記グループを特定するためのシステム識別子を保持し、

前記デバイスセルは、前記システム識別子に対応する CPU セルのリストから構成されるシステム構成情報を保持し、

前記デバイスセルは、前記コマンドと共に前記 CPU セルから送信された前記システム識別子と保持した前記システム構成情報から該 CPU セルが所属するグループの中から任意の一つの CPU セルを選択し、該選択した CPU セルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、

CPU セルは、前記デバイスセルから前記応答メッセージを受信すると、該応答メッセージにしたがって前記デバイスセルの処理結果を取得する請求項 1 乃至 6 のいずれか 1 項記載のマルチプロセッサシステム。

【請求項 8】 少なくとも 1 つの CPU を備えた複数の CPU セルを有し、前記 CPU セルが複数のグループに分割され、該グループ毎に異なるオペレーティングシステムで動作するマルチプロセッサシステムでデバイスを共用するためのデバイス共有方法であって、

前記 CPU セルとネットワークを介して接続された前記デバイスを含むデバイスセルに、前記デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備えておき、

前記デバイスセルで、前記 CPU セルからコマンドを受信すると、該コマンドの発行元に対応するデバイス管理情報を検索し、前記コマンドにより更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させるデバイス共

有方法。

【請求項 9】 前記デバイス管理情報は、

前記複数の CPUセルにそれぞれ対応して設けられた請求項 8 記載のデバイス共有方法。

【請求項 10】 前記デバイス管理情報は、

前記グループの数、及び前記 CPUセルの数に一致しない、任意の数だけ設けられた請求項 8 記載のデバイス共有方法。

【請求項 11】 前記 CPUセルで、自機が利用可能な前記デバイスの情報をテーブル形式で保持し、

システム立ち上げ時に該 CPUセルに割り当てられたデバイスを優先的に利用する請求項 8 記載のデバイス共有方法。

【請求項 12】 前記デバイスセルに複数の同じデバイスを備え、

前記デバイスセルは、前記デバイス管理情報で指定された処理を、前記複数のデバイスのうちの任意のデバイスに実行させる請求項 8 乃至 11 のいずれか 1 項記載のデバイス共有方法。

【請求項 13】 前記 CPUセルに、前記 CPUから発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、

前記デバイスセルに、前記コマンドを分解して複数の命令を抽出するコマンド解析部を備え、抽出された複数の命令により更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させる請求項 8 乃至 12 のいずれか 1 項記載のデバイス共有方法。

【請求項 14】 前記 CPUセルで、自機が所属する前記グループを特定するためのシステム識別子を保持し、

前記デバイスセルで、前記システム識別子に対応する CPUセルのリストから構成されるシステム構成情報を保持し、

前記デバイスセルは、前記コマンドと共に前記 CPUセルから送信された前記システム識別子と保持した前記システム構成情報から該 CPUセルが所属するグループの中から任意の一つの CPUセルを選択し、該選択した CPUセルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、

CPUセルは、前記デバイスセルから前記応答メッセージを受信すると、該応答メッセージにしたがって前記デバイスセルの処理結果を取得する請求項8乃至13のいずれか1項記載のデバイス共有方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数のCPUを有するマルチプロセッサシステムに関し、特に複数のOSで動作する環境下でデバイスを各OSで共用するマルチプロセッサシステム及びデバイス共有方法に関する。

【0002】

【従来の技術】

(第1従来例)

図14はマルチプロセッサシステムの第1従来例の構成を示すブロック図である。なお、図14は特許文献1で開示された構成である。

【0003】

図14に示すように、第1従来例のマルチプロセッサシステムは、複数のCPUセル200と複数のデバイスセル206とを有し、それらがネットワーク205を介して接続された構成である。

【0004】

CPUセル200は、複数のCPU201と、CPU201の動作を制御する制御回路202と、プログラムやデータを記憶するメモリ203と、ネットワーク205との通信を制御する通信回路204とを備えている。

【0005】

デバイスセル206は、外部記憶装置や入出力装置のように各CPUセル200の処理で利用する複数のデバイス209と、デバイス209に対するインタフェース部であるIO制御回路208と、ネットワーク205との通信を制御する通信回路207とを備えている。CPUセル200及びデバイスセル206は、それぞれが有する通信回路204、207を利用し、コマンドやデータを含むパケットをネットワーク205を介して送受信する。なお、図14では、CPUセ

ル 200 及びデバイスセル 206 をそれぞれ複数 (図 14 では 2 つ) 備えているため、各 CPU セルを 200-1、200-2 のように符号を付与することで区別し、各デバイスセルを 206-1、206-2 のように符号を付与することで区別している。また、CPU セル内の複数の CPU (図 14 では 2 つ) を 201-11、201-12 のように符号を付与することで区別し、制御回路、メモリ、及び通信回路には、CPU セルに付与した符号に対応して、例えば制御回路 202-1、メモリ 203-1、及び通信回路 204-1 のように符号を付与している。また、デバイスセル内の複数のデバイス (図 14 では 2 つ) を 209-11、209-12 のように符号を付与することで区別し、IO 制御回路、及び通信回路には、デバイスセルに付与した符号に対応して、例えば IO 制御回路 208-1、及び通信回路 207-1 のように符号を付与している。

【0006】

このようなマルチプロセッサシステムでは、システムの立ち上げ時に、CPU セル 200 の通信回路 204 に当該 CPU セル 200 で使用するデバイスセル 206 の宛先情報を設定し、デバイスセル 206 の通信装置 207 に当該デバイスセル 207 を使用する CPU セル 200 の宛先情報を設定しておく。

【0007】

そして、システムの運用が開始され、CPU 201 からコマンドが発行されると、該コマンドは通信回路 204 へ渡され、システム立ち上げ時に設定された宛先情報にしたがって通信回路 204 からネットワーク 205 を経由して該当するデバイスセル 206 へ該コマンドが送信される。

【0008】

デバイスセル 206 では、CPU セル 200 から送信されたコマンドを通信回路 207 で受信し、その内容に基づいて自セル内のデバイス 209 を制御する。そして、当該コマンドに対する応答メッセージをシステム立ち上げ時に設定された宛先情報にしたがって通信回路 207 からネットワーク 205 を経由して該当する CPU セル 200 へ送信する。CPU セル 200 では、送信された応答メッセージを通信回路 204 で受信し、CPU 201 へ伝達する。

【0009】

” ”
(第2従来例)

第2従来例のマルチプロセッサシステムは、図14に示した構成を利用して複数のCPUセルが備えるメモリを共有化するための技術である。このような技術は、例えば特許文献2に開示されている。

【0010】

第2従来例のマルチプロセッサシステムでは、例えば、CPUセル200-1のCPU201-11からCPUセル200-2のメモリ203-2に対してアクセス命令が発行されると、該アクセス命令は通信回路204-1によりネットワーク205を経由してCPUセル200-2へ送信される。

【0011】

CPUセル200-2は、メモリ203-2に対するアクセス命令を通信回路204-2で受信すると、該アクセス命令にしたがってメモリ203-2にアクセスする。そして、当該アクセス命令に対する応答メッセージを通信回路204-2からネットワーク205を経由してCPUセル200-1へ送信する。CPUセル200-1はCPUセル200-2から送信された応答メッセージを通信回路204-1で受信し、CPU201-11へ伝達する。

【0012】

このような技術を第1従来例の構成に適用することで、CPUセル200-1に割り当てられたデバイスセル206-1であってもCPUセル200-2からアクセスすることが可能になる。

【0013】

例えば、CPUセル200-2のCPU201-21から入出力命令が発行されると、該入出力命令は通信回路204-2へ渡され、通信回路204-2からネットワーク205を経由してCPUセル200-1へ送信される。

【0014】

CPUセル200-1は、CPUセル200-2から送信された入出力命令を通信回路204-1で受信すると、該入出力命令をシステム立ち上げ時に設定された宛先情報にしたがって通信回路204-1によりネットワーク205を経由して該当するデバイスセル206-1へ送信する。

【0015】

デバイスセル 206-1 は、CPUセル 200-1 から送信されたコマンドを通信回路 207-1 で受信し、その内容に基づいて自セル内のデバイス 209 を制御する。そして、システム立ち上げ時に設定された宛先情報にしたがって当該コマンドに対する応答メッセージを通信回路 207-1 からネットワーク 205 を経由して該当する CPUセル 200-1 へ送信する。CPUセル 200-1 は、送信された応答メッセージを通信回路 204-1 で受信し、ネットワーク 205 を経由して CPUセル 200-2 へ配信する。CPUセル 200-2 は応答メッセージを通信回路 204-2 で受信し、CPU 201-21 へ伝達する。

【0016】

(第3従来例)

第3従来例のマルチプロセッサシステムは、異なるオペレーティングシステム(OS)で動作している CPUセル間でストレージデバイス(ハードディスク装置や光ディスク装置等の記録装置)を共有する手法であり、例えば、非特許文献 1 及び特許文献 3 に開示された技術である。

【0017】

図 15 はマルチプロセッサシステムの第3従来例の構成を示すブロック図である。

【0018】

図 15 に示すように、第3従来例のマルチプロセッサシステムは、図 14 に示した第1従来例のマルチプロセッサシステムに加えて、複数のストレージデバイスを備えたストレージセル 212 を有し、ストレージセル 212 がファイバチャネル(以下、FCと略す)ネットワーク 211 により複数のデバイスセル 206 に接続された構成である。

【0019】

デバイスセル 206 には、FC ネットワーク 211 を介して通信するための FC デバイス 210 を備えている。

【0020】

また、ストレージセル 212 は、データを蓄積する複数のストレージデバイス

215と、ストレージデバイス215の動作を制御するストレージ制御ブリッジ214と、FCネットワーク211を介してデバイスセル206との通信を制御するホスト入出力手段213とを備えている。なお、図15では、図14に示したマルチプロセッサシステムと同様に、各CPUセルを200-1、200-2のように符号を付与することで区別し、各デバイスセルを206-1、206-2のように符号を付与することで区別している。また、CPUセル内の複数のCPU（図15では2つ）を201-11、201-12のように符号を付与することで区別し、制御回路、メモリ、及び通信回路には、CPUセルに付与した符号に対応して、例えば制御回路202-1、メモリ203-1、及び通信回路204-1のように符号を付与している。また、デバイスセル内のデバイス、FCデバイス、IO制御回路、及び通信回路には、デバイスセルに付与した符号に対応して、例えばデバイス206-1、FCデバイス210-1、IO制御回路208-1、及び通信回路207-1のように符号を付与している。また、ストレージセル212内の複数のストレージデバイス（図15では2つ）を215-1、215-2のように符号を付与することで区別している。

【0021】

このようなマルチプロセッサシステムでは、各CPUセル200からストレージセル212にアクセスする場合に、対応するデバイスセル206内のFCデバイス210を経由することでストレージセル212にアクセスすることが可能になる。したがって、異なるOSで動作する複数のCPUセル200であってもストレージセル212を共用できる。

【0022】

【特許文献1】

特開2002-229967号

【特許文献2】

特開2000-259596号

【特許文献3】

特開2000-347815号

【0023】

【非特許文献 1】

ファイバーチャネル協議会編、「ファイバーチャネル技術解説書」、論創社

【0024】**【発明が解決しようとする課題】**

上記したような従来のマルチプロセッサシステムのうち、第 1 従来例のマルチプロセッサシステムでは、デバイスセルに対してシステム起動時に割り当てられた CPU セルからしかアクセスできないという問題がある。

【0025】

例えば、図 14 に示したマルチプロセッサシステムにおいて、CPU セル 200-1 に割り当てられたデバイス 209-11 を CPU セル 200-2 で利用するためには、CPU セル 200-1 及び CPU セル 200-2 の動作を一旦停止し、CPU セル 200-1、CPU セル 200-2、及びデバイスセル 206-1 の各通信回路に設定される宛先情報を変更した後、CPU セル 200-1 及び CPU セル 200-2 を再起動する必要がある。

【0026】

したがって、デバイスセル 206 に対するアクセスの切り替えのためにシステムを停止させる必要があるので、マルチプロセッサシステムの性能、可用性が低下するという問題があった。

【0027】

第 2 従来例のマルチプロセッサシステムでは、システム起動時に割り当てられていないデバイスセルに対してアクセスする場合、ネットワークを 2 回経由してコマンドを送信する必要があるため、入出力命令等のレイテンシが増加し、システム性能が悪化する問題がある。また、CPU セルで次の入出力命令を発行するためには、以前に発行した入出力命令が、宛先となるデバイスセルに割り当てられた CPU セルへ到達したことが保証された後になるため、命令がネットワーク内を 1 往復する時間のレイテンシがターンアラウンドタイムになり、入出力命令に対する処理性能が悪化する問題もある。

【0028】

さらに、第 1 従来例と第 2 従来例とを組み合わせることで、1 つのデバイスを

複数のCPUセルで利用することが可能になるが、これはCPUセルが1つのOSで動作していることが前提となる。

【0029】

第3従来例のマルチプロセッサシステムでは、CPUセル及びデバイスセル間の通信機能を提供しているネットワークとは別に、ストレージデバイスを共有するためのFCネットワークが必要となるため、ハードウェアのコストが増大する問題がある。また、CPUセルからストレージデバイスへアクセスする際にFCデバイスを中継するため、ストレージデバイス用のデバイスドライバとは別にFCデバイス用のデバイスドライバを開発する必要がある。

【0030】

第3従来例のマルチプロセッサシステムは、複数のFCデバイスを1つのデバイスドライバで管理することで信頼性の向上や負荷分散による性能の向上が可能とされている。しかしながら、例えば、1つのマルチプロセッサシステムがパーティショニングされて4つのOSで動作するようなケースでは、信頼性を確保するために1つのOSあたり2つのFCデバイスが必要となる。そのため、パーティショニング時のハードウェアコストが増大するという問題がある。

【0031】

また、複数のFCデバイスを1つのデバイスドライバが管理する場合、特許文献3には、そのFCデバイスに対する入出力命令のレイテンシ及びTATを改善するために、複数のFCデバイスを最適に選択するための方法や技術が開示されていない。

【0032】

さらに、デバイスドライバは、通常、管理するFCデバイスへのアクセスが競合しないように排他制御を行うため、例えば2つのアプリケーションプログラムが同時に同じFCデバイスに対するアクセス要求を発行した場合に、排他制御により片方のアプリケーションのアクセス要求がブロックされて性能が低下するという問題があった。

【0033】

本発明は上記したような従来技術の有する問題点を解決するためになされた

ものであり、複数のOSで動作する環境下であってもデバイスを共有化することが可能な簡易な構成のマルチプロセッサシステムを提供することを第1の目的とする。

【0034】

また、デバイスに対する入出力命令等のレイテンシやターンアラウンドタイムを短縮すると共に、デバイスに対するアクセス要求の競合やネットワークのトラフィック量の増大を軽減し、デバイスに対するアクセス性能を向上させたマルチプロセッサシステムを提供することを第2の目的とする。

【0035】

さらに、ハードウェアのコストを増大させることなく、デバイスの信頼性の向上や負荷分散による性能の向上が可能なマルチプロセッサシステムを提供することを第3の目的とする。

【0036】

【課題を解決するための手段】

上記目的を達成するため本発明のマルチプロセッサシステムは、少なくとも1つのCPUを備えた複数のCPUセルを有し、前記CPUセルが複数のグループに分割され、該グループ毎に異なるオペレーティングシステムで動作するマルチプロセッサシステムであって、

前記複数のCPUセル間で共用されるデバイス、

並びに前記デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備え、前記CPUセルからコマンドを受信すると、該コマンドの発行元に対応するデバイス管理情報を検索し、前記コマンドにより更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させるデバイス管理部を備えた、前記CPUセルとネットワークを介して接続されたデバイスセルを有する構成である。

【0037】

このとき、前記デバイス管理情報は、

前記複数のCPUセルにそれぞれ対応して設けられていてもよく、
前記グループの数、及び前記CPUセルの数に一致しない、任意の数だけ設けら

れていてもよい。

【0038】

また、前記CPUセルは、利用可能な前記デバイスの情報をテーブル形式で保持し、

システム立ち上げ時に自機に割り当てられたデバイスを優先的に利用してもよい。

【0039】

さらに、前記デバイスセルは、

複数の同じデバイスを備え、

前記デバイス管理部は、

前記デバイス管理情報で指定された処理を、前記複数のデバイスのうちの任意のデバイスに実行させてもよく、

前記CPUセルは、

前記CPUから発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、

前記デバイスセルは、

前記コマンドを分解して複数の命令を抽出するコマンド解析部を備え、

前記デバイス管理部は、抽出された複数の命令により更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させてもよい。

前記CPUセルは、自機が所属する前記グループを特定するためのシステム識別子を保持し、

前記デバイスセルは、前記システム識別子に対応するCPUセルのリストから構成されるシステム構成情報を保持し、

前記デバイスセルは、前記コマンドと共に前記CPUセルから送信された前記システム識別子と保持した前記システム構成情報から該CPUセルが所属するグループの中から任意の一つのCPUセルを選択し、該選択したCPUセルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、

CPUセルは、前記デバイスセルから前記応答メッセージを受信すると、該応答メッセージにしたがって前記デバイスセルの処理結果を取得してもよい。

【0040】

一方、本発明のデバイス共有方法は、少なくとも1つのCPUを備えた複数のCPUセルを有し、前記CPUセルが複数のグループに分割され、該グループ毎に異なるオペレーティングシステムで動作するマルチプロセッサシステムでデバイスを共用するためのデバイス共有方法であって、

前記CPUセルとネットワークを介して接続された前記デバイスを含むデバイスセルに、前記デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備えておき、

前記デバイスセルで、前記CPUセルからコマンドを受信すると、該コマンドの発行元に対応するデバイス管理情報を検索し、前記コマンドにより更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させる方法である。

【0041】

このとき、前記デバイス管理情報は、

前記複数のCPUセルにそれぞれ対応して設けられていてもよく、
前記グループの数、及び前記CPUセルの数に一致しない、任意の数だけ設けられていてもよい。

【0042】

また、前記CPUセルで、自機が利用可能な前記デバイスの情報をテーブル形式で保持し、

システム立ち上げ時に該CPUセルに割り当てられたデバイスを優先的に利用してもよい。

【0043】

さらに、前記デバイスセルに複数の同じデバイスを備え、

前記デバイスセルは、前記デバイス管理情報で指定された処理を、前記複数のデバイスのうちの任意のデバイスに実行させてもよく、

前記CPUセルに、前記CPUから発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、

前記デバイスセルに、前記コマンドを分解して複数の命令を抽出するコマンド

解析部を備え、抽出された複数の命令により更新される前記デバイス管理情報で指定された処理を前記デバイスに実行させてもよい。

【0044】

また、前記CPUセルで、自機が所属する前記グループを特定するためのシステム識別子を保持し、

前記デバイスセルで、前記システム識別子に対応するCPUセルのリストから構成されるシステム構成情報を保持し、

前記デバイスセルは、前記コマンドと共に前記CPUセルから送信された前記システム識別子と保持した前記システム構成情報から該CPUセルが所属するグループの中から任意の一つのCPUセルを選択し、該選択したCPUセルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、

CPUセルは、前記デバイスセルから前記応答メッセージを受信すると、該応答メッセージにしたがって前記デバイスセルの処理結果を取得してもよい。

【0045】

上記のようなマルチプロセッサシステム及びデバイス共有方法では、デバイスセルに、デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備え、CPUセルからコマンドを受信した場合に、該コマンドの発行元に対応するデバイス管理情報を検索し、コマンドにより更新されるデバイス管理情報で指定された処理をデバイスに実行させることで、各CPUセルまたは各グループから同じデバイスにアクセスすることが可能になるため、簡素なハードウェアによりデバイスの共有化が可能になる。

【0046】

また、CPUセルで、自機が利用可能な前記デバイスの情報をテーブル形式で保持し、システム立ち上げ時に該CPUセルに割り当てられたデバイスを優先的に利用することで、例えば、デバイスに対する入出力命令のレイテンシはネットワークを1往復する時間となる。また、ターンアラウンドタイムは、CPUセル内で閉じた時間となるため、ネットワークを1往復するデバイスへのアクセス時間と比較して短くできる。また、デバイスセルに複数の同じデバイスを備え、デバイス管理情報で指定された処理を複数のデバイスのうちの任意のデバイスに

実行させることで、例えば、任意のデバイスを利用している最中に障害が発生した場合、他のデバイスを利用して処理を続行できる。

【0047】

さらに、CPUセルに、CPUから発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、デバイスセルに、コマンドを分解して複数の命令を抽出するコマンド解析部を備え、抽出された複数の命令により更新されるデバイス管理情報で指定された処理をデバイスに実行させることで、デバイスに対する入出力命令等によるネットワークのトラフィックが低減する。

【0048】

また、CPUセルで、自機が所属する前記グループを特定するためのシステム識別子を保持し、デバイスセルで、システム識別子に対応するCPUセルのリストから構成されるシステム構成情報を保持し、デバイスセルは、コマンドと共にCPUセルから送信されたシステム識別子と保持したシステム構成情報から該CPUセルが所属するグループの中から任意の一つのCPUセルを選択し、該選択したCPUセルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、CPUセルは、デバイスセルから応答メッセージを受信すると、該応答メッセージにしたがってデバイスセルの処理結果を取得することで、デバイスセルでコマンド処理中にシステム構成が変更された場合でもコマンドを発行したグループに正しく応答メッセージを返すことができる。

【0049】

【発明の実施の形態】

次に本発明について図面を参照して説明する。

【0050】

（第1の実施の形態）

図1は本発明のマルチプロセッサシステムの第1の実施の形態の構成を示すブロック図であり、図2は図1に示したネットワーク上で伝送されるパケットの構成を示す模式図である。また、図3は図1に示したデバイスのポート構成を示すブロック図であり、図4は図1に示したデバイス制御ブリッジが備えるデバイス管理情報の構成を示すブロック図である。

【0051】

図1に示すように、第1の実施の形態のマルチプロセッサシステムは、複数のCPUセル11と複数のデバイスセル12とを有し、それらがネットワーク13を介して接続された構成である。

【0052】

CPUセル11は、複数のCPU14と、CPU14の動作を制御するCPU制御ブリッジ16と、プログラムやデータを記憶するメモリ15と、ネットワーク13との通信を制御する通信回路17とを備えている。

【0053】

CPU14はメモリ15に格納されたOSの管理下でアプリケーションプログラムを実行することにより所定の処理を実行する。このとき、入出力命令やメモリアクセス命令等のコマンドをCPU制御ブリッジ16へ出力する。CPU14から発行されるコマンドが入出力命令の場合は対象となるデバイス（入力装置または出力装置）のポートが指定され、メモリアクセス命令の場合はアクセスするメモリのアドレスが指定される。

【0054】

デバイスセル12は、デバイス18と、デバイス18に対するインタフェース部であるデバイス制御ブリッジ19と、ネットワーク13との通信を制御する通信回路20とを備えている。

【0055】

デバイス制御ブリッジ19は、デバイス管理部21と、CPUセル11毎に用意される記憶領域である複数のデバイス管理情報22とを備えている。デバイス管理情報22は、対応するCPUセル11から見える（アクセス可能な）デバイスの状態を保持するためのものである。

【0056】

CPUセル11及びデバイスセル12は、それぞれが有する通信回路を利用して、コマンドやデータを含むパケットをネットワーク13を経由して送受信する。

。

【0057】

CPUセル11及びデバイスセル12が備える各通信回路には、それらを識別するための一意なID（このIDはCPUセル11やデバイスセル12を特定するIDでもある）が付与される。また、各CPU14には、それらを識別するための一意なCPUIDが付与される。

【0058】

図2に示すように、パケットにはCPU14から発行された入出力命令やメモリアクセス命令と共にCPUIDを含む情報がデータ領域25に格納される。また、送付先ID23には送付先のCPUセル11またはデバイスセル12が備える通信回路のIDが格納され、送付元ID24には送付元のCPUセル11またはデバイスセル12が備える通信回路のIDが格納される。ネットワーク13はパケット中の送付先ID23及び送付元ID24を参照し、送付元の通信回路から送付先IDで特定される通信回路に対してパケットを配信する。

【0059】

なお、図1では、3つのCPUセル11と2つのデバイスセル12を有するマルチプロセッサシステムの構成例を示し、各CPUセルを11-1、11-2、11-3のように符号を付与することで区別し、各デバイスセルを12-1、12-2のように符号を付与することで区別している。また、CPUセル内の複数のCPU（図1では2つ）は14-11、14-12のように符号を付与することで区別し、CPU制御ブリッジ、メモリ、及び通信回路は、CPUセルに付与した符号に対応して、例えばCPU制御ブリッジ16-1、メモリ15-1、及び通信回路17-1のように符号が付与される。また、デバイスセル内のデバイス、デバイス制御ブリッジ、及び通信回路は、デバイスセルに付与した符号に対応して、例えばデバイス18-1、デバイス制御ブリッジ19-1、及び通信回路20-1のように符号が付与される。

【0060】

図1に示すマルチプロセッサシステムは、システム起動時に、CPUセル11-1及びCPUセル11-2からなるシステムAと、CPUセル11-3からなるシステムBとにパーティショニングされ、システムAとシステムBとが異なるOSで動作するものとする。

【0061】

ここで、デバイス18は、図3に示すように、動作モードを設定するためのモード設定ポート26と、処理を設定するための処理設定ポート27と、処理の結果を格納する結果読出ポート28とを備えているものとする。デバイス18は、CPUセル11からのコマンドによって処理設定ポート27に書き込みが行われ、モード設定ポート26及び処理設定ポート27に書き込まれた内容に基づいて処理を実行し、その処理結果を結果読出ポート28に書き込む。このデバイス18には、第1のモードで実行する第1の処理及び第2の処理、または第2のモードで実行する第1の処理及び第2の処理の合わせて4つの処理が存在する。

【0062】

デバイス管理情報22は、図4に示すように、設定された動作モードを一時的に保持するモード設定レジスタ29と、処理結果を一時的に保持する結果レジスタ30とを備えている。このモード設定レジスタ29及び結果レジスタ30を有するデバイス管理情報22は、図1に示すように、各CPUセル11に対応して設けられている。図1に示すように、デバイス管理情報には、デバイス18に付与した符号及び対応するCPUセルに付与された符号に対応して、例えばデバイス18-1とCPU11-3に対応するデバイス管理情報には22-13という符号が付与される。また、図4に示すように、モード設定レジスタ及び結果レジスタにはデバイス管理情報に付与した符号に対応して、例えばモード設定レジスタ29-11及び結果レジスタ30-11のように符号を付与する。

【0063】

次に、図1に示す2つのシステムA、Bで1つのデバイス18-1を共用する場合を例にして、本実施形態のマルチプロセッサシステムの動作について説明する。なお、システムの立ち上げ時、CPUセル11-1の通信回路17-1には、システムAで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が書き込まれ、CPUセル11-3の通信回路17-3には、システムBで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が書き込まれ、CPUセル11-2の通信回路17-2には、システムAで利用するデバイス18-1に対応する

入出力ポートの宛先としてCPUセル11-1の宛先情報が書き込まれるものとする。

【0064】

また、システムAに属するCPUセル11-2のCPU14-22は、デバイス18-1に対して、モード設定ポート26-1に第1のモードを書き込み、処理設定ポート27-1に第1の処理を書き込み、結果読出しポート28-1からその結果を読み出すものとする。また、システムBに属するCPUセル11-3のCPU14-31は、デバイス18-1に対して、モード設定ポート26-1に第2のモードを書き込み、処理設定ポート27-1に第2の処理を書き込み、結果読出しポート28-1からその結果を読み出すものとする。また、ここでは、システムAとシステムBとが同時にデバイス18-1をアクセスする場合を説明する。

【0065】

まず、CPUセル11-2のCPU14-22がデバイス18-1のモード設定ポート26-1に第1のモードを書き込む入出力命令を発行すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-2によりネットワーク13を経由してCPUセル11-1へ該入出力命令を含むパケットが送信される。このとき送信されるパケットの送付先ID23はCPUセル11-1を示し、送付元ID24はCPUセル11-2を示し、データ領域25にはCPU14-22で発行された入出力命令とCPU14-22を示すCPUIDが格納される。

【0066】

CPUセル11-1は、CPUセル11-2から送信されたパケットを通信回路17-1で受信すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-1からネットワーク13を経由して該当するデバイスセル12-1へ当該パケットを送信する。この時、パケットの送付先ID23はデバイスセル12-1を示し、送付元ID24はCPU11-1を示し、データ領域25にはCPU14-22から発行された設定入出力命令、そのCPUID、及びCPUセル11-2を示す要求元IDが格納される。

【0067】

デバイスセル 12-1 は、CPUセル 11-1 から送信されたパケットを通信回路 20-1 で受信すると、該パケットから抽出した入出力命令をデバイス管理部 21-1 へ転送する。デバイス管理部 21-1 は、パケット送付元の CPUセル 11-1 に対応するデバイス管理情報 22-11 を参照し、受け取った入出力命令及びデバイス管理情報 22-11 の内容に基づき、必要に応じてデバイス 18-1 の制御やデバイス管理情報 22-11 の更新を行う。この場合、第 1 のモードを書き込む入出力命令であるため、デバイス管理情報 22-11 のモードレジスタ 29-11 に第 1 のモードを書き込み、処理を終了する。

【0068】

次に、CPUセル 11-3 の CPU14-31 がデバイス 18-1 のモード設定ポート 26-1 に第 2 のモードを書き込む入出力命令を発行すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路 17-3 によりネットワーク 13 を経由して該当するデバイスセル 12-1 へ該入出力命令を含むパケットが送信される。この時、パケットの送付先 ID 23 はデバイスセル 12-1 を示し、送付元 ID 24 は CPUセル 11-3 を示し、データ領域 25 には CPU14-31 から発行された入出力命令とその CPU ID が格納される。

【0069】

デバイスセル 12-1 は、CPUセル 11-3 から送信されたパケットを通信回路 20-1 で受信すると、該パケットから抽出された入出力命令をデバイス管理部 21-1 へ転送する。デバイス管理部 21-1 は、パケット送付元の CPUセル 11-3 に対応するデバイス管理情報 22-13 を参照し、受け取った入出力命令及びデバイス管理情報 22-13 の内容に基づき、必要に応じてデバイス 18-1 の制御やデバイス管理情報 22-13 の更新を行う。この場合、第 2 のモードを書き込む入出力命令であるため、デバイス管理情報 22-13 のモードレジスタ 29-13 に第 2 のモードを書き込み、処理を終了する。

【0070】

次に、CPUセル 11-2 の CPU14-22 がデバイス 18-1 の処理設定ポート 27-1 に第 1 の処理を書き込む入出力命令を発行すると、第 1 のモード

を書き込む入出力命令の発行時と同様の手順で、デバイスセル 12-1 のデバイス管理部 21-1 に当該入出力命令が伝達される。

【0071】

デバイス管理部 21-1 は、パケット送信元の CPU セル 11-1 に対応するデバイス管理情報 22-11 のモードレジスタ 29-11 を読出し、デバイス 18-1 のモード設定ポート 26-1 にその内容（第 1 のモード）を書き込む。続いて、第 1 の処理を書き込む入出力命令にしたがって処理設定ポート 27-1 に第 1 の処理を書き込む。

【0072】

デバイス 18-1 は、処理設定ポート 27-1 への書き込みを受けて、第 1 のモードの第 1 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から当該処理結果を読出し、その内容を結果レジスタ 30-11 に書き込む。

【0073】

次に、CPU セル 11-3 の CPU 14-22 がデバイス 18-1 の処理設定ポート 27-1 に第 2 の処理を書き込む入出力命令を発行すると、第 2 のモードを書き込む入出力命令の発行時と同様の手順で、デバイスセル 12-1 のデバイス管理部 21-1 に当該入出力命令が伝達される。

【0074】

デバイス管理部 21-1 は、送信元 CPU セル 11-3 に対応するデバイス管理情報 22-13 のモードレジスタ 29-13 を読出し、デバイス 18-1 のモード設定ポート 26-1 にその内容（第 2 のモード）を書き込む。続いて、第 2 の処理を書き込む入出力命令にしたがって処理設定ポート 27-1 に第 2 の処理を書き込む。

【0075】

デバイス 18-1 は、処理設定ポート 27-1 への書き込みを受けて、第 2 のモードの第 2 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き

込まれたことを確認すると、結果読出ポート 28-1 から当該処理結果を読み出し、その内容を結果レジスタ 30-13 に書き込む。

【0076】

次に、CPUセル 11-2 の CPU14-22 がデバイス 18-1 の結果読出ポート 28-1 からの読出し入出力命令を発行すると、第 1 のモードを書き込む入出力命令の発行時と同様の手順で、デバイスセル 12-1 のデバイス管理部 21-1 に当該入出力命令が伝達される。デバイス管理部 21-1 は、パケット送信元の CPUセル 11-1 に対応するデバイス管理情報 22-11 の結果レジスタ 29-11 を読み出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。通信回路 20-1 は、送付元の CPUセル 11-1 に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先 ID 23 は CPUセル 11-1 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 には応答メッセージと CPU14-22 を示す CPU ID と要求元 CPUセル 11-2 を示す要求元 ID が格納される。

【0077】

CPUセル 11-1 は、この応答メッセージを含むパケットを通信回路 17-1 で受信すると、ネットワーク 13 を経由して CPUセル 11-2 に配信する。この時、パケットの送付先 ID 23 は CPUセル 11-2 を示し、送付元 ID 24 は CPUセル 11-1 を示し、データ領域 25 には応答メッセージと CPU14-22 を示す CPU ID が格納される。CPUセル 11-2 は、この応答メッセージを含むパケットを通信回路 17-2 で受信すると、該パケットから応答メッセージを抽出し、CPU14-22 へ伝達する。

【0078】

次に、CPUセル 11-3 の CPU14-31 がデバイス 18-1 の結果読出ポート 28-1 からの読出し入出力命令を発行すると、第 2 のモードを書き込む入出力命令の発行時と同様の手順で、デバイスセル 12-1 のデバイス管理部 21-1 に当該入出力命令が伝達される。デバイス管理部 21-1 は、送信元の CPUセル 11-3 に対応するデバイス管理情報 22-13 の結果レジスタ 29-13 を読み出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。

通信回路 20-1 は、送付元の CPU セル 11-3 に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先 ID 23 は CPU セル 11-3 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 には応答メッセージと CPU 14-31 を示す CPU ID が格納される。CPU セル 11-3 は、この応答メッセージを含むパケットを通信回路 17-3 で受信すると、該パケットから応答メッセージを抽出し、CPU 14-31 へ伝達する。

【0079】

以上のような手順で動作することで、システム A に属する CPU セル 11-2 の CPU 14-22 は、デバイス 18-1 に対してモード設定ポート 26-1 に第 1 のモードを書き込み、処理設定ポート 27-1 に第 1 の処理を書き込み、結果読出しポート 28-1 からその処理結果を読み出すことができる。また、システム B に属する CPU セル 11-3 の CPU 14-31 は、デバイス 18-1 に対してモード設定ポート 26-1 に第 2 のモードを書き込み、処理設定ポート 27-1 に第 2 の処理を書き込み、結果読出しポート 28-1 からその処理結果を読み出すことができる。

【0080】

なお、本実施形態のマルチプロセッサシステムでは、例えば、デバイスセル 12-2 にデバイスセル 12-1 が有するデバイス 18-1 と同じデバイスを設けてもよい。その場合、システム立ち上げ時、例えば、CPU セル 11-1 の通信回路 17-1 には、システム A で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-2 に対応する入出力ポートの宛先としてデバイスセル 12-2 の宛先情報を設定する。また、CPU セル 11-3 の通信回路 17-3 には、システム B で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-2 に対応する入出力ポートの宛先としてデバイスセル 12-2 の宛先情報を設定する。また、CPU セル 11-2 の通信回路 17-2 には、システム A で利用するデバイス 18-1 及びデバイス 18-2 に対応する入出力ポートの宛先として CPU セル 11-1 の宛先情報を設定する。

このような構成では、システム A 及びシステム B は、デバイス 18-1 とデバイス 18-2 とが利用可能なデバイスとして見える。

【0081】

したがって、各システムのデバイスドライバは、2 つ存在するデバイス 18-1 とデバイス 18-2 とを利用することで、負荷を分散させてシステムの性能を向上させることが可能になる。また、フェイルオーバーを行うことでデバイスの信頼性を向上させることもできる。しかも、それに必要なデバイス数は 2 個であり、マルチプロセッサシステムで動作する OS の数に比例して増加することがない。

【0082】

さらに、本実施形態のマルチプロセッサシステムでは、例えば、CPU セル 11-1 と CPU セル 11-2 の 2 つの CPU セルで構成されるシステム A において、デバイス 18-1 を 2 つの異なるデバイス 18-11 とデバイス 18-12 に見せることもできる。システム立ち上げ時に、CPU セル 11-1 の通信回路 17-1 には、システム A で利用する仮想的なデバイス 18-11 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-12 に対応する入出力ポートの宛先として CPU セル 11-2 を設定する。また、CPU セル 11-2 の通信回路 17-2 には、システム A で利用するデバイス 18-11 に対応する入出力ポートの宛先として CPU セル 11-1 の宛先情報を設定し、デバイス 18-12 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定する。このような構成では、システム A は、デバイス 18-11 とデバイス 18-12 との 2 個のデバイスが利用可能なデバイスとして見える。

【0083】

したがって、システム A では、デバイス 18-1 を管理するデバイスドライバが 2 つのアプリケーションプログラムから同時にアクセス要求を受けても、デバイス 18-11 とデバイス 18-12 とに使い分けることができる。したがって、同一デバイスにアクセスするための排他制御を行う必要がなくなり、システム性能を向上させることができる。

【0084】

よって、第1の実施の形態のマルチプロセッサシステムは、各OSのデバイスドライバを改変することなくデバイスを共有化することができる。

【0085】

また、マルチプロセッサシステムで動作する複数のOS間で、簡素なハードウェアによりデバイスの共有化が可能であり、ハードウェアのコストを低減できる。さらに、デバイスの信頼性向上や負荷分散のために必要なデバイス数は、マルチプロセッサシステムで動作するOSの数に依存せずに一定であるため、ハードウェアのコストを低減できる。また、実際には1つしかないデバイスを、仮想的に複数のデバイスに見せることが可能となり、デバイスドライバが2つのアプリケーションプログラムから同時にアクセス要求を受けても、複数の仮想的なデバイスを使い分けることができるようになる。したがって、同一デバイスにアクセスするための排他制御を行う必要が無くなり、システム性能を向上させることができる。

【0086】

(第2の実施の形態)

第1の実施の形態では、デバイスセル12に、各CPUセル11に対応するデバイス管理情報22を備えることで異なるOSで動作する複数のシステムで1つのデバイスを共用できるようにしていた。したがって、CPUセル11が多数存在する場合は、デバイス管理情報22も多数必要となるため、ハードウェアコストが増大するおそれがある。また、仮想的に見せることができるデバイスの数は、システムを構成するCPUセル11の数に限定されるため、システムを構成するCPUセルの数が1つの場合は仮想的に2つのデバイスに見せることができなかった。本実施形態のマルチプロセッサシステムは、CPUセル11とデバイス管理情報22を1対1で対応させない例である。

【0087】

図5は本発明の第2の実施の形態のマルチプロセッサシステムが有するデバイス管理情報の構成を示すブロック図である。

【0088】

図5に示すように、第2の実施の形態のマルチプロセッサシステムは、第1の実施の形態で備えていたレジスタに加えて、デバイス管理情報22にCPUセルレジスタ31、入出力ポートベースレジスタ32、及び入出力ポート長レジスタ33を備えた構成である。これらのレジスタの値はシステム立ち上げ時にそれぞれ設定される。

【0089】

例えば、図1に示したシステムAとシステムBとを有するマルチプロセッサシステムでデバイス18-1を共有する場合、デバイス管理情報22をシステムAとシステムBに対応して2つ設けておく。そして、システム立ち上げ時、CPUセル11-1の通信回路17-1にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報を設定し、CPUセル11-3の通信回路17-3にはシステムBで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報を設定し、CPUセル11-2の通信回路17-2にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてCPUセル11-1の宛先情報を設定する。

【0090】

また、システムAに対応するデバイス管理情報22のCPUセルレジスタ31にCPUセル11-1を示すIDを格納し、入出力ポートベースレジスタ32にシステムAで設定される入出力ポートのベースアドレスを格納し、入出力ポート長レジスタ33にその範囲を設定する。

【0091】

さらに、システムBに対応するデバイス管理情報22のCPUセルレジスタ31にCPUセル11-3を示すIDを格納し、入出力ポートベースレジスタ32にシステムBで設定される入出力ポートのベースアドレスを格納し、入出力ポート長レジスタ33にその範囲を設定する。

【0092】

デバイス管理部21は、あるシステム（CPUセル11）を送信元とする入出力命令を受信すると、対応するデバイス管理情報22を検索し、CPUセルレジ

スタ 31 の内容が送信元 ID と一致し、かつ入出力ポートの番号が入出力ポートベースレジスタ 32 と入出力ポート長レジスタ 33 で示される範囲（入出力ベースレジスタ 32 以上、入出力ベースレジスタ 32 + 入出力ポート長レジスタ 33 未満）を検出する。そして、検出したデバイス管理情報 22 を利用して第 1 の実施の形態と同様にデバイス 18 を制御する。

【0093】

このように、デバイス管理情報 22 に CPU セルレジスタ 31、入出力ポートベースレジスタ 32、及び入出力ポート長レジスタ 33 を備え、入出力命令を受信したときに該当するデバイス管理情報 22 を検索することで、CPU セル 11 に対してデバイス管理情報 22 が 1 対 1 に対応しない構成も可能になる。すなわち、CPU セル 11 の数に依存しない、任意の数のデバイス管理情報 22 を用意する構成が可能となり、そのデバイス管理情報 22 と CPU セル 11 の割り当ても自由に設定することができる。これにより、CPU セルの数に関係なく、例えば 2 つのデバイス管理情報 22 を用意すれば 2 つのシステムでデバイス 18 を共用できるため、ハードウェアコストを低減できる。

【0094】

また、例えば、2 つのデバイス管理情報 22 を同じ CPU セル 11 に割り当て、異なる入出力ポートを設定すれば、システムを構成する CPU セル 11 の数が 1 つでも 2 つのデバイスが存在するように見せることができる。

【0095】

（第 3 の実施の形態）

第 1 及び第 2 の実施の形態では、1 つの OS に対して複数の同じデバイス 18 を割り当てた場合にデバイス 18 をどのように使い分けるかは特に提案していなかった。そのため、利用するデバイス 18 が別の CPU セル 11 に割り当てられている場合、従来と同様に、そのデバイス 18 に対する入出力命令のレイテンシ及びターンアラウンドタイムが悪化する問題は解決できていなかった。

【0096】

図 6 は本発明のマルチプロセッサシステムの第 3 の実施の形態の処理手順を示すフローチャートであり、図 7 は第 3 の実施の形態のマルチプロセッサシステム

の各システムが備えるデバイスドライバが保持する情報を示すテーブル図である。なお、図6は、1つのOSに対して複数の同じデバイスを割り当てた場合のデバイスドライバによる使い分けの手順を示している。デバイスドライバは、利用可能なデバイスの情報をテーブル形式で保持しており、図7はそのテーブルの各エントリの構成を示している。

【0097】

図7に示すCPUセルID34は対応するデバイス18に割り当てられているCPUセルのIDを示す。また、入出力ポートベース35は対応するデバイス18が備える入出力ポートのベースアドレスが格納される。これらCPUセルID34及び入出力ポートベース35の情報はシステム立ち上げ時に設定される。ロックビット36はデバイス18が利用中であるか否かを示す。

【0098】

図6に示すように、任意のアプリケーションからデバイス18に対するアクセス要求を受付けると、デバイスドライバは、ステップ101にてテーブルを検索し、当該デバイスドライバが動作しているCPU14が含まれるCPUセルのIDとCPUセルID34の内容が一致し、かつロックビット36が「0」のエントリを1つ検出する。

【0099】

続いて、ステップ102にて該当するエントリが検出されたか否かを判定し、エントリが検出された場合はステップ105に移行する。

【0100】

一方、ステップ102にて該当するエントリが検出されなかった場合は、ステップ103にて再度テーブルを検索し、ロックビット36が「0」のエントリを1つ検出する。そして、ステップ104にて該当するエントリが検出されたか否かを再度判定し、該当するエントリが検出された場合はステップ105に移行する。また、検出されなかった場合はステップ101に戻る。

ステップ105では、検出したエントリを利用するデバイス18として選択し、そのエントリのロックビット36を「1」に設定する。次に、ステップ106にて検出したエントリの入出力ポートベース35の内容に基づいてアクセスするデ

バイス 18 の入出力ポート番号を計算し、デバイス 18 にアクセスする。

【0101】

デバイス 18 へのアクセスが完了すると、ステップ 107 にて検出したエントリのロックビット 36 を「0」に設定する。

【0102】

以上の手順で処理することで、デバイスドライバは同一 CPU セル 11 の通信回路 17 に割り当てられたデバイス 18 を優先的に利用することになる。この場合、デバイス 18 への入出力命令のレイテンシはネットワーク 13 を 1 往復する時間となる。また、ターンアラウンドタイムは、CPU セル 11 内で閉じた時間となるため、ネットワーク 13 を 1 往復するデバイス 18 へのアクセス時間と比較して短くできる。

【0103】

したがって、第 3 の実施の形態のマルチプロセッサシステムによれば、1 つの OS に複数の同じデバイスを割り当てた場合に、デバイスに対する入出力命令のレイテンシ及びターンアラウンドタイムを短縮できる。

【0104】

(第 4 の実施の形態)

第 4 の実施の形態のマルチプロセッサシステムは、第 1 ～ 第 3 の実施の形態の変形例として、デバイスセル 12 に複数のデバイス 18 を有する構成である。その場合、デバイスセル 12 には、各デバイス 18 に対応するデバイス管理部 21 及びデバイス管理情報 22 が用意される。各デバイス 18 へのアクセスは対応するデバイス管理部 21 を経由して行われる。

【0105】

図 8 は第 4 の実施の形態のマルチプロセッサシステムが有するデバイスセルの構成を示すブロック図である。

【0106】

図 8 に示すように、本実施形態のマルチプロセッサシステムは、デバイスセル 12-1 にデバイス 18-1a 及びデバイス 18-1b を備え、デバイス制御ブリッジ 19-1 にデバイス 18-1a 及びデバイス 18-1b がそれぞれ接続さ

れた構成である。

【0107】

デバイス制御ブリッジ19-1は、デバイス18-1aに対応するデバイス管理部21-1a及びデバイス管理情報22-11a, 12a, 13aと、デバイス18-1bに対応するデバイス管理部21-1b及びデバイス管理情報22-11b, 12b, 13bとを備えている。

【0108】

このような構成において、図8に示すデバイスセル12-1の通信回路20-1でデバイス18-1aに対する入出力命令を受信すると、デバイス管理部21-1aを用いてデバイス18-1aが制御される。また、デバイス18-1bに対する入出力命令を受信すると、デバイス管理部21-1bを用いてデバイス18-1bが制御される。

【0109】

なお、本実施形態のマルチプロセッサシステムでは、デバイスセル12に同種のデバイスを複数備え、それらを1つのデバイス管理部21で制御することも可能である。その場合のデバイス管理部21の構成を図9に示す。

【0110】

図9は第4の実施の形態のマルチプロセッサシステムが有するデバイスセルの変形例の構成を示すブロック図である。

【0111】

図9に示すデバイスセル12-1は、同種のデバイス18-1a及びデバイス18-1bを備え、デバイス制御ブリッジ19-1にデバイス18-1a及びデバイス18-1bがそれぞれ接続された構成である。

【0112】

デバイス制御ブリッジ19-1は、デバイス18-1a及びデバイス18-1bそれぞれに対応するデバイス管理部21-1a及びデバイス管理情報22-11a, 12a, 13aを備えている。

【0113】

デバイス管理部21は、デバイス18-1a及びデバイス18-1bの利用状

況をそれぞれ監視し、負荷分散を行う。また、一方のデバイスが障害等で利用できなくなった場合は、他方のデバイスを利用することによりフェイルオーバーを実現し、信頼性を向上させる。

【0114】

図10は図9に示したデバイス管理部が保持する情報を示すテーブル図である。

【0115】

図10に示すように、本実施形態のデバイス管理部21-1では、デバイス18-1aに対応するロックビット37-1a及びデバイス18-1bに対応するロックビット37-1bを保持している。これらのロックビット37-1a、37-1bはそれぞれデバイス18-1a及びデバイス18-1bの利用可否状況を示している。例えば、デバイス18-1aが利用可能な場合はロックビット37-1aが「0」となり、現在利用中あるいは障害により利用できない場合は「1」となる。

【0116】

次に、第4の実施の形態のマルチプロセッサシステムのデバイス管理部の動作について図11を用いて説明する。

【0117】

図11は本発明のマルチプロセッサシステムの第4の実施の形態の処理手順を示すフローチャートである。なお、図11はデバイス管理部21-1が通信回路20-1を介して入出力命令を受信し、その入出力命令に対応する処理を行って実際にデバイスにアクセスする必要が生じた場合の動作を示している。

【0118】

図11に示すように、デバイス管理部21-1は、通信回路20-1を介して入出力命令を受信すると、ステップ111にて利用可能なデバイスを見出すために値が「0」のロックビット37-1を検索する。

【0119】

次に、ステップ112にて値が「0」のロックビット37-1を検出したか否かを判定し、「0」のロックビット37-1を検出した場合は（以降の説明では

ロックビット 37-1 a が「0」であったとして説明を行う) ステップ 113 に移行する。また、検出していない場合はステップ 111 に戻る。

【0120】

次に、ステップ 113 にて検出したロックビット 37-1 a に対応するデバイス 18-1 a を利用するデバイスとして選択し、該ロックビット 37-1 a に「1」を書き込む。そして、ステップ 113 で選択したデバイス 18-1 a にアクセスし、処理の完了を待ち受ける。

【0121】

次に、ステップ 115 にてデバイス 18-1 a で障害が発生したか否かを判定し、処理の完了を待っている間にデバイス 18-1 a の障害が検出された場合はステップ 110 に戻る。また、正常にデバイス 18-1 a の処理が完了した場合はステップ 116 に移行する。ステップ 116 では、ロックビット 37-1 a に「0」を書き込む。

【0122】

以上のような手順で処理することで、デバイス管理部 21-1 は、例えば、ある入出力命令にしたがってデバイス 18-1 a にアクセスしている最中に、他の入出力命令を受けてデバイス 18-1 にアクセスする必要が生じた場合でも、デバイス 18-1 b を利用することができる。これにより、デバイス 18-1 に対するアクセスの負荷をデバイス 18-1 a 及びデバイス 18-1 b を利用して分散できる。

【0123】

また、例えば、デバイス 18-1 a を利用している最中に障害が発生した場合、もう 1 つのデバイス 18-1 b を利用して処理を続行できる。これにより、障害時のフェイルオーバーを実現できる。

【0124】

したがって、第 4 の実施の形態のマルチプロセッサシステムでは、デバイスセルに複数のデバイスを接続することが可能になる。また、同種のデバイスを接続し、1 つのデバイス管理部 21 で管理することによりデバイスドライバの改変を行うことなく、負荷分散及びフェイルオーバーを実現できる。

(第5の実施の形態)

図12は本発明のマルチプロセッサシステムの第5の実施の形態の構成を示すブロック図であり、図13は図12に示したデバイス制御ブリッジが備えるデバイス管理情報の構成を示すブロック図である。

【0125】

図12に示すように、第5の実施の形態のマルチプロセッサシステムは、CPUセル11にコマンド送出回路38を備え、デバイスセル12のデバイス管理部21にコマンド解析部39を備えた点が第1の実施の形態と異なっている。

【0126】

CPUセル11は、コマンド送出回路38によりCPU14から発行された複数の入出力命令をまとめたコマンドを生成し、通信回路17によりパケット化してネットワーク13へ送出する。デバイスセル12は、通信回路20によりネットワーク13からパケットを受信すると、該パケットから抽出したコマンドをデバイス管理部21へ転送する。

【0127】

デバイス管理部21は、入出力命令ではないコマンドを受信した場合、コマンド解析部39に該コマンドを渡して解析を依頼する。コマンド解析部39は該コマンドを複数の入出力命令に分解してデバイス管理部21に返送する。デバイス管理部21はコマンド解析部39から返送された複数の入出力命令にしたがって処理を実行する。

【0128】

ここで、デバイス18には、第1の実施の形態と同様に、動作モードの種類を設定するためのモード設定ポート26と、処理の種類を設定するための処理設定ポート27と、処理の結果を格納する結果読出ポート28とを備えている。また、CPU14は、モード設定ポート26への書き込みアクセス及び処理設定ポート27への書き込みアクセスを同時に生成するものとする。この場合、デバイス管理情報22は、図13に示すように、結果レジスタ30のみで構成できる。

【0129】

次に、図12に示す2つのシステムA、Bで1つのデバイス18-1を共有す

る場合を例にして、本実施形態のマルチプロセッサシステムの動作について説明する。なお、システムの立ち上げ時、CPUセル11-1の通信回路17-1にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が設定され、CPUセル11-3の通信回路17-3にはシステムBで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が設定され、CPUセル11-2の通信回路17-2にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてCPUセル11-1の宛先情報が設定されるものとする。

【0130】

また、システムAに属するCPUセル11-2のCPU14-22は、デバイス18-1に対して、モード設定ポート26-1に第1のモードを書き込み、処理設定ポート27-1に第1の処理を書き込み、結果読出しポート28-1からその結果を読み出すとする。また、システムBに属するCPUセル11-3のCPU14-31は、デバイス18-1に対して、モード設定ポート26-1に第2のモードを書き込み、処理設定ポート27-1に第2の処理を書き込み、結果読出しポート28-1からその結果を読み出すとする。ここでは、システムAとシステムBとが同時にデバイス18-1にアクセスする場合で説明する。

【0131】

まず、CPUセル11-2のCPU14-22が第1のモード及び第1の処理に設定するための入出力命令をまとめたコマンドを生成すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-2によりネットワーク13を経由してCPUセル11-1へ該コマンドを含むパケットが送信される。このとき送信されるパケットの送付先ID23はCPUセル11-1を示し、送付元ID24はCPUセル11-2を示し、データ領域25にはCPUセル11-2で発行されたコマンドとCPU14-22を示すCPUIDが格納される。

【0132】

CPUセル11-1は、CPUセル11-2から送信されたパケットを通信回路17-1で受信すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-1からネットワーク13を経由して該当するデバイスセル12

ー 1 へ当該パケットを送信する。この時、パケットの送付先 ID 23 はデバイスセル 12-1 を示し、送付元 ID 24 は CPU 11-1 を示し、データ領域 25 には CPU セル 11-2 で発行されたコマンド、及び CPU セル 11-2 を示す要求元 ID が格納される。

【0133】

デバイスセル 12-1 は、このコマンドを含むパケットを通信回路 20-1 で受信すると、該パケットからコマンドを抽出し、デバイス管理部 21-1 に伝達する。デバイス管理部 21-1 は、受信したコマンドをコマンド解析部 39-1 で解析し、モード設定ポート 26-1 に対する第 1 のモードへの設定入出力命令、及び処理設定ポート 27-1 に対する第 1 の処理への設定入出力命令を取得する。そして、デバイス 18-1 のモード設定ポート 26-1 を第 1 のモードに設定し、処理設定ポート 27-1 を第 1 の処理に設定する。

【0134】

デバイス 18-1 は、処理設定ポート 27-1 に対する書き込みを受けて、第 1 のモードの第 1 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。

【0135】

デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から処理結果を読み出し、その内容をパケットの送付元 ID が示す CPU セル 11-1 に対応する結果レジスタ 30-11 へ書き込む。

【0136】

次に、CPU セル 11-3 の CPU 14-31 が第 2 のモード及び第 2 の処理に設定するための入出力命令をまとめたコマンドを生成すると、CPU セル 11-3 はコマンド送出回路 38-3 を利用して該コマンドをデバイスセル 12-1 に送出する。この時、CPU セル 11-3 の通信回路 17-3 から出力されるパケットの送付先 ID 23 はデバイスセル 18-1 を示し、送付元 ID 24 は CPU セル 11-3 を示し、データ領域 25 には該コマンドが格納される。

【0137】

デバイスセル 12-1 は、このコマンドを含むパケットを通信回路 20-1 で受信すると、該パケットからコマンドを抽出し、デバイス管理部 21-1 に伝達する。デバイス管理部 21-1 は、受信したコマンドをコマンド解析部 39-1 で解析し、モード設定ポート 26-1 に対する第 2 のモードへの設定入出力命令、及び処理設定ポート 27-1 に対する第 2 の処理への設定入出力命令を取得する。そして、デバイス 18-1 のモード設定ポート 26-1 を第 2 のモードに設定し、処理設定ポート 27-1 を第 2 の処理に設定する。

【0138】

デバイス 18-1 は、処理設定ポート 27-1 に対する書き込みを受けて、第 2 のモードの第 2 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。

【0139】

デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から処理結果を読み出し、その内容をパケットの送付元 ID が示す CPU セル 11-3 に対応する結果レジスタ 30-13 へ書き込む。

【0140】

次に、CPU セル 11-2 の CPU 14-22 がデバイス 18-1 の結果読出ポート 28-1 への読出し入出力命令を発行すると、CPU セル 11-2 は通信回路 17-2 により該入出力命令を含むパケットを生成し、ネットワーク 13 を経由して CPU セル 11-1 に送信する。この時、CPU セル 11-2 の通信回路 17-2 から出力されるパケットの送付先 ID 23 は CPU 11-1 を示し、送付元 ID 24 は CPU セル 11-2 を示し、データ領域 25 には該入出力命令と CPU 14-22 を示す CPU ID が格納される。

【0141】

CPU セル 11-1 は、このコマンドを含むパケットを通信回路 17-1 で受信すると、システム立ち上げ時に設定された宛先情報にしたがって CPU セル 11-2 から受信した入出力命令を含むパケットをネットワーク 13 を経由して該当するデバイスセル 12-1 へ送信する。この時、パケットの送付先 ID 23 は

デバイスセル 12-1 を示し、送付元 ID 24 は CPU 11-1 を示し、データ領域 25 には入出力命令と CPU 14-22 を示す CPU ID と CPU セル 11-2 を示す要求元 ID が格納される。

【0142】

デバイスセル 12-1 は、この入出力命令を含むパケットを通信回路 20-1 で受信すると、該パケットから入出力命令を抽出し、デバイス管理部 21-1 に伝達する。デバイス管理部 21-1 は、送信元の CPU セル 11-1 に対応するデバイス管理情報 22-11 の結果レジスタ 29-11 を読出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。通信回路 20-1 は、送付元の CPU セル 11-1 に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先 ID 23 は CPU セル 11-1 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 部には、応答メッセージと CPU 14-22 を示す CPU ID と要求元の CPU セル 11-2 を示す要求元 ID が格納される。

【0143】

CPU セル 11-1 は、デバイスセル 12-1 から応答メッセージを含むパケットを通信回路 17-1 で受信すると、ネットワーク 13 を経由して CPU セル 11-2 へ配信する。この時、パケットの送付先 ID 23 は CPU セル 11-2 を示し、送付元 ID 24 は CPU セル 11-1 を示し、データ領域 25 には応答メッセージと CPU 14-22 を示す CPU ID が格納される。

【0144】

CPU セル 11-2 は、この応答メッセージを含むパケットを通信回路 17-2 で受信すると、該パケットから応答メッセージを抽出し、CPU 14-22 へ伝達する。

【0145】

次に、CPU セル 11-3 の CPU 14-31 がデバイス 18-1 の結果読出ポート 28-1 への読出し入出力命令を発行すると、CPU セル 11-3 は通信回路 17-3 により該入出力命令を含むパケットを生成し、システム立ち上げ時に設定された宛先情報にしたがってネットワーク 13 を経由してデバイスセル 1

2-1へ送信する。この時、CPUセル11-3の通信回路17-3から出力されるパケットの送付先ID23はデバイスセル12-1を示し、送付元ID24はCPUセル11-3を示し、データ領域25には該入出力命令とCPU14-31を示すCPUIDが格納される。

【0146】

デバイスセル12-1は、この入出力命令を含むパケットを通信回路20-1で受信すると、該パケットから入出力命令を抽出し、デバイス管理部21-1に伝達する。デバイス管理部21-1は、送信元のCPUセル11-3に対応するデバイス管理情報22-13の結果レジスタ29-13を読み出し、その内容を応答メッセージとして通信回路20-1に伝達する。通信回路20-1は、送付元のCPUセル11-3に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先ID23はCPUセル11-3を示し、送付元ID24はデバイスセル12-1を示し、データ領域25部には、応答メッセージとCPU14-31を示すCPUIDが格納される。

【0147】

CPUセル11-3は、この応答メッセージを含むパケットを通信回路17-3で受信すると、該パケットから応答メッセージを抽出し、CPU14-31へ伝達する。

【0148】

以上のような手順で動作することで、システムAに属するCPUセル11-2のCPU14-22は、デバイス18-1に対してモード設定ポート26-1に第1のモードを書き込み、処理設定ポート27-1に第1の処理を書き込み、結果読出しポート28-1からその処理結果を読み出すことができる。また、システムBに属するCPUセル11-3のCPU14-31は、デバイス18-1に対してモード設定ポート26-1に第2のモードを書き込み、処理設定ポート27-1に第2の処理を書き込み、結果読出しポート28-1からその処理結果を読み出すことができる。

【0149】

なお、本実施形態のマルチプロセッサシステムでは、例えば、デバイスセル1

2-2 にデバイスセル 12-1 が有するデバイス 18-1 と同じデバイスを設けてもよい。その場合、システム立ち上げ時、例えば、CPUセル 11-1 の通信回路 17-1 にはシステム A で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-2 に対応する入出力ポートの宛先としてデバイスセル 12-2 の宛先情報を設定する。また、CPUセル 11-3 の通信回路 17-3 にはシステム B で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-2 に対応する入出力ポートの宛先としてデバイスセル 12-2 の宛先情報を設定する。また、CPUセル 11-2 の通信回路 17-2 にはシステム A で利用するデバイス 18-1 及びデバイス 18-2 に対応する入出力ポートの宛先として CPUセル 11-1 の宛先情報を設定する。このような構成では、システム A 及びシステム B は、デバイス 18-1 とデバイス 18-2 とが利用可能なデバイスとして見える。

【0150】

したがって、各システムのデバイスドライバは、2つ存在するデバイス 18-1 とデバイス 18-2 とを利用することで、負荷を分散させてシステムの性能を向上させることが可能になる。また、フェイルオーバーを行うことでデバイスの信頼性を向上させることもできる。しかも、それに必要なデバイス数は2個であり、マルチプロセッサシステムで動作する OS の数に比例して増加することがない。

【0151】

さらに、本実施形態のマルチプロセッサシステムでは、例えば、CPUセル 11-1 と CPUセル 11-2 の二つの CPUセルで構成されるシステム A において、デバイス 18-1 を2つの異なるデバイス 18-11 とデバイス 18-12 に見せることもできる。システム立ち上げ時に、CPUセル 11-1 の通信回路 17-1 には、システム A で利用する仮想的なデバイス 18-11 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定し、デバイス 18-12 に対応する入出力ポートの宛先として CPUセル 11-2 を設定する。また、CPUセル 11-2 の通信回路 17-2 にはシステム A で利用するデバイ

ス 18-11 に対応する入出力ポートの宛先として CPU セル 11-1 の宛先情報を設定し、デバイス 18-12 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報を設定する。このような構成では、システム A はデバイス 18-11 とデバイス 18-12 との 2 個のデバイスが利用可能なデバイスとして見える。

【0152】

したがって、システム A では、デバイス 18-1 を管理するデバイスドライバが 2 つのアプリケーションプログラムから同時にアクセス要求を受けても、デバイス 18-11 とデバイス 18-12 とに使い分けることができる。したがって、同一デバイスにアクセスするための排他制御を行う必要が無くなり、システム性能を向上させることができる。

【0153】

よって、第 5 の実施の形態のマルチプロセッサシステムは、各 OS のデバイスドライバを改変することなくデバイスを共有化することができる。

【0154】

また、マルチプロセッサシステムで動作する複数の OS 間で、簡素なハードウェアによりデバイスの共有化が可能であり、ハードウェアのコストを低減できる。さらに、デバイスの信頼性向上や負荷分散のために必要なデバイス数は、マルチプロセッサシステムで動作する OS の数に依存せずに一定であるため、ハードウェアのコストを低減できる。また、実際には 1 つしかないデバイスを、仮想的に複数のデバイスに見せることが可能となり、デバイスドライバが 2 つのアプリケーションプログラムから同時にアクセス要求を受けても、複数の仮想的なデバイスを使い分けることができるようになる。したがって、同一デバイスにアクセスするための排他制御を行う必要が無くなり、システム性能を向上させることができる。

【0155】

なお、第 5 の実施の形態で示した構成（CPU セル 11 にコマンド送出回路 38 を備え、デバイスセル 12 のデバイス管理部 21 にコマンド解析部 39 を備えた構成）は、上述した第 2 ～第 4 の実施の形態にも適応可能である。

【0156】

(第6の実施の形態)

第5の実施の形態では、デバイス18-1が備えるモード設定ポート26-1及び処理設定ポート27-1に対する入出力命令を1つのコマンドにまとめて送信する例を示した。第6の実施の形態は、それに結果読出しポート28-1に対する入出力命令を加える例である。この場合、デバイス管理情報22には何の情報も持つ必要がない。マルチプロセッサシステムの構成は図12で示す第5の実施の形態と同様であるため、その説明は省略する。

【0157】

次に、図12に示す2つのシステムA、Bで1つのデバイス18-1を共有する場合を例にして、本実施形態のマルチプロセッサシステムの動作について説明する。なお、システムの立ち上げ時、CPUセル11-1の通信回路17-1にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が設定され、CPUセル11-3の通信回路17-3にはシステムBで利用するデバイス18-1に対応する入出力ポートの宛先としてデバイスセル12-1の宛先情報が設定され、CPUセル11-2の通信回路17-2にはシステムAで利用するデバイス18-1に対応する入出力ポートの宛先としてCPUセル11-1の宛先情報が設定されるものとする。

【0158】

また、システムAに属するCPUセル11-1のCPU14-12は、デバイス18-1に対して、モード設定ポート26-1に第1のモードを書き込み、処理設定ポート27-1に第1の処理を書き込み、結果読出しポート28-1からその結果を読み出すとする。また、システムBに属するCPUセル11-3のCPU14-31は、デバイス18-1に対して、モード設定ポート26-1に第2のモードを書き込み、処理設定ポート27-1に第2の処理を書き込み、結果読出しポート28-1からその結果を読み出すとする。ここでは、システムAとシステムBとが同時にデバイス18-1にアクセスする場合で説明する。

【0159】

まず、CPUセル11-2のCPU14-22が第1のモード及び第1の処理

に設定するための入出力命令と処理結果を格納するメモリ 15 上のアドレスとをまとめたコマンドを生成すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路 17-2 によりネットワーク 13 を経由して CPU セル 11-1 へ該コマンドを含むパケットが送信される。このとき送信されるパケットの送付先 ID 23 は CPU セル 11-1 を示し、送付元 ID 24 は CPU セル 11-2 を示し、データ領域 25 には CPU セル 11-2 で発行されたコマンドと CPU 14-22 を示す CPU ID が格納される。

【0160】

CPU セル 11-1 は、CPU セル 11-2 から送信されたパケットを通信回路 17-1 で受信すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路 17-1 からネットワーク 13 を経由して該当するデバイスセル 12-1 へ当該パケットを送信する。この時、パケットの送付先 ID 23 はデバイスセル 12-1 を示し、送付元 ID 24 は CPU 11-1 を示し、データ領域 25 には CPU セル 11-2 で発行されたコマンド、及び CPU セル 11-2 を示す要求元 ID が格納される。

【0161】

デバイスセル 12-1 は、このコマンドを含むパケットを通信回路 20-1 で受信すると、該パケットからコマンドを抽出し、デバイス管理部 21-1 に伝達する。デバイス管理部 21-1 は、受信したコマンドをコマンド解析部 39-1 で解析し、モード設定ポート 26-1 に対する第 1 のモードへの設定入出力命令、処理設定ポート 27-1 に対する第 1 の処理への設定入出力命令、及び結果読出しポート 28-1 からの読出し入出力命令を取得する。そして、デバイス 18-1 のモード設定ポート 26-1 を第 1 のモードに設定し、処理設定ポート 27-1 を第 1 の処理に設定する。

【0162】

デバイス 18-1 は、処理設定ポート 27-1 に対する書き込みを受けて、第 1 のモードの第 1 の処理を実行し、処理結果を結果読出しポート 28-1 に書き込む。

【0163】

デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から処理結果を読み出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。通信回路 20-1 は、送付元の CPU セル 11-1 に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先 ID 23 は CPU セル 11-1 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 部には、応答メッセージと受信したパケットに含まれていたアドレスが格納される。

【0164】

CPU セル 11-1 は、この応答メッセージを含むパケットを通信回路 17-1 で受信すると、該パケットから応答メッセージを抽出し、指定されたアドレスに対して応答メッセージを書き込むメモリアクセス命令として処理する。

【0165】

次に、CPU セル 11-3 の CPU 14-31 が第 2 のモード及び第 2 の処理に設定するための入出力命令と処理結果を格納するメモリ 15 上のアドレスとをまとめたコマンドを生成すると、CPU セル 11-3 はコマンド送出回路 38-1 を利用して該コマンドをデバイスセル 12-1 に送出する。この時、CPU セル 11-1 の通信回路 17-1 から出力されるパケットの送付先 ID 23 はデバイスセル 18-1 を示し、送付元 ID 24 は CPU セル 11-3 を示し、データ領域 25 には該コマンドが格納される。

【0166】

デバイスセル 12-1 は、このコマンドを含むパケットを通信回路 20-1 で受信すると、該パケットからコマンドを抽出し、デバイス管理部 21-1 に伝達する。デバイス管理部 21-1 は、受信したコマンドをコマンド解析部 39-1 で解析し、モード設定ポート 26-1 に対する第 2 のモードへの設定入出力命令、処理設定ポート 27-1 に対する第 2 の処理への設定入出力命令、及び結果読出しポート 28-1 からの読出し入出力命令を取得する。そして、デバイス 18-1 のモード設定ポート 26-1 を第 2 のモードに設定し、処理設定ポート 27-1 を第 2 の処理に設定する。

【0167】

デバイス 18-1 は、処理設定ポート 27-1 に対する書き込みを受けて、第 2 のモードの第 2 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。

【0168】

デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から処理結果を読み出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。通信回路 20-1 は、送付元の CPU セル 11-3 に対してこの応答メッセージを含むパケットを送信する。この時、パケットの送付先 ID 23 は CPU セル 11-3 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 部には、応答メッセージと受信したパケットに含まれていたアドレスが格納される。

【0169】

CPU セル 11-3 は、この応答メッセージを含むパケットを通信回路 17-3 で受信すると、該パケットから応答メッセージを抽出し、指定されたアドレスに対して応答メッセージを書き込むメモリアクセス命令として処理する。

【0170】

この状態で、CPU 14-22 及び CPU 14-31 は、指定したメモリ 15 上のアドレスに処理結果が書き込まれたことを確認し、その内容を読み出して処理結果を取得する。

【0171】

以上のような手順で動作することで、システム A に属する CPU セル 11-2 の CPU 14-22 は、デバイス 18-1 に対してモード設定ポート 26-1 に第 1 のモードを書き込み、処理設定ポート 27-1 に第 1 の処理を書き込み、その処理結果を指定したメモリ 15 上のアドレスから読み出すことができる。また、システム B に属する CPU セル 11-3 の CPU 14-31 は、デバイス 18-1 に対してモード設定ポート 26-1 に第 2 のモードを書き込み、その処理結果を指定したメモリ 15 上のアドレスから読み出すことができる。

【0172】

したがって、第 6 の実施の形態のマルチプロセッサシステムは、第 5 の実施の

形態で得られた効果に加えて、入出力命令によるネットワーク 13 のトラフィックを削減する効果、及びデバイス管理情報 22 を簡素化する効果が得られる。

【0173】

(第 7 の実施の形態)

第 6 の実施の形態では、デバイスセル 12-1 でパケットを受信すると、送付元の CPU セル 11-1 に対して応答メッセージを送信していた。第 7 の実施の形態では、デバイスセル 12-1 が応答メッセージを含むパケットの送付先を選択する例である。このため、本実施形態の各 CPU セル 11 の通信回路 17 は、所属するシステムを特定するためのシステム識別子を保持する。また、デバイスセル 12 の通信回路 20 は、システム識別子に対応する CPU セル 11 のリストを保持する。

【0174】

図 12 に示す例では、システム立ち上げ時、CPU セル 11-1 の通信回路 17-1 にはシステム A で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報が設定されると共にシステム A を示すシステム識別子が格納される。また、CPU セル 11-3 の通信回路 17-3 にはシステム B で利用するデバイス 18-1 に対応する入出力ポートの宛先としてデバイスセル 12-1 の宛先情報が設定されると共にシステム B を示すシステム識別子が格納される。マルチプロセッサシステムの構成は図 12 で示す第 6 の実施の形態と同様であるため、その説明は省略する。

【0175】

次に、図 12 に示す 2 つのシステム A、B で 1 つのデバイス 18-1 を共有する場合を例にして、本実施形態のマルチプロセッサシステムの動作について説明する。

【0176】

なお、ここではシステム A に属する CPU セル 11-2 の CPU 14-22 が、デバイス 18-1 に対して、モード設定ポート 26-1 に第 1 のモードを書き込み、処理設定ポート 27-1 に第 1 の処理を書き込み、結果読出しポート 28-1 からその結果を読み出すとする。

【0177】

CPUセル11-2のCPU14-22が第1のモード及び第1の処理に設定するための入出力命令と処理結果を格納するメモリ15上のアドレスとをまとめたコマンドを生成すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-2によりネットワーク13を経由してCPUセル11-1へ該コマンドを含むパケットが送信される。このとき送信されるパケットの送付先ID23はCPUセル11-1を示し、送付元ID24はCPUセル11-2を示し、データ領域25にはCPUセル11-2で発行されたコマンドとCPU14-22を示すCPUIDが格納される。

【0178】

CPUセル11-1は、CPUセル11-2から送信されたパケットを通信回路17-1で受信すると、システム立ち上げ時に設定された宛先情報にしたがって通信回路17-1からネットワーク13を経由して該当するデバイスセル12-1へ当該パケットを送信する。この時、パケットの送付先ID23はデバイスセル12-1を示し、送付元ID24はCPU11-1を示し、データ領域25にはCPUセル11-2で発行されたコマンド、CPUセル11-2を示す要求元ID、及びシステム立ち上げ時に設定されたシステムAを示すシステム識別子が格納される。

【0179】

デバイスセル12-1は、このコマンドを含むパケットを通信回路20-1で受信すると、該パケットからコマンドを抽出し、デバイス管理部21-1に伝達する。デバイス管理部21-1は、受信したコマンドをコマンド解析部39-1で解析し、モード設定ポート26-1に対する第1のモードへの設定入出力命令、処理設定ポート27-1に対する第1の処理への設定入出力命令、及び結果読出しポート28-1からの読出し入出力命令を取得する。そして、デバイス18-1のモード設定ポート26-1を第1のモードに設定し、処理設定ポート27-1を第1の処理に設定する。

【0180】

デバイス18-1は、処理設定ポート27-1に対する書き込みを受けて、第

1 のモードの第 1 の処理を実行し、処理結果を結果読出ポート 28-1 に書き込む。

【0181】

デバイス管理部 21-1 は、結果読出ポート 28-1 を監視し、処理結果が書き込まれたことを確認すると、結果読出ポート 28-1 から処理結果を読み出し、その内容を応答メッセージとして通信回路 20-1 に伝達する。

【0182】

この段階で、システム A が CPU セル 11-2 のみで構成されるように変更され、通信回路 20-1 のシステム構成情報が変更されていたとする。

【0183】

通信回路 20-1 は、受信したパケットに含まれていたシステム識別子と、保持するシステム構成情報から、そのシステムを構成する CPU セル 11の中から任意の 1 つを選択し、その CPU セル 11 に向けて応答メッセージを含むパケットを送出する。ここでは、CPU セル 11-2 のみでシステム A が構成されているため、CPU セル 11-2 に対して応答メッセージを含むパケットを送信する。このとき、パケットの送付先 ID 23 は CPU セル 11-2 を示し、送付元 ID 24 はデバイスセル 12-1 を示し、データ領域 25 には応答メッセージと受信したパケットに含まれていたアドレスが格納される。

CPU セル 11-2 は、デバイスセル 12-1 からパケットを受信すると、指定されたアドレスに対して応答メッセージを書き込むメモリアクセス命令として処理する。そして、指定したメモリ 15 上のアドレスに処理結果が書き込まれたことを確認し、その内容を読み出して処理結果を取得する。

【0184】

以上のような手順で動作することで、システム構成が変更された場合でもコマンドを発行したシステムに正しく応答メッセージを返すことができる。したがって、デバイスセルでコマンド処理中であってもシステム構成を変更することが可能になる。

【0185】

また、本実施形態のマルチプロセッサシステムでは、デバイスセル 12-1 の

通信回路 20-1 で応答メッセージを返送する CPU セル 11 を選択する際に、コマンドを発行した CPU セルと同一システム内の CPU セルをラウンドロビン、あるいはランダムに選択することでネットワーク 13 や CPU セルの負荷を分散することも可能である。

【0186】

第 7 の実施の形態は、第 6 の実施の形態で得られた効果に加えて、デバイス 18 で処理を行っている最中にシステム構成が変更されても正しく応答を返すことができる。また、ネットワーク 13 や CPU セル 11 の負荷を分散することができる。

【0187】

なお、上述した第 1 ～第 7 の実施の形態の変形例として、デバイス管理部 21 にプロセッサを備え、上記デバイス管理部 21 で実現した機能の一部をソフトウェアプログラムで動作する該プロセッサで実現してもよい。

【0188】

【発明の効果】

本発明は以上説明したように構成されているので、以下に記載する効果を奏する。

【0189】

デバイスセルに、デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備え、CPU セルからコマンドを受信した場合に、該コマンドの発行元に対応するデバイス管理情報を検索し、コマンドにより更新されるデバイス管理情報で指定された処理をデバイスに実行させることで、各 CPU セルまたは各グループから同じデバイスにアクセスすることが可能になるため、簡素なハードウェアによりデバイスの共有化が可能になり、ハードウェアのコストを低減できる。さらに、デバイスの信頼性向上や負荷分散のために必要なデバイス数は、マルチプロセッサシステムで動作する OS の数に依存せずに一定であるため、ハードウェアのコストを低減できる。

【0190】

また、CPU セルで、自機が利用可能な前記デバイスの情報をテーブル形式で

保持し、システム立ち上げ時に該CPUセルに割り当てられたデバイスを優先的に利用することで、例えば、デバイスに対する入出力命令のレイテンシはネットワークを1往復する時間となる。また、ターンアラウンドタイムは、CPUセル内で閉じた時間となるため、ネットワークを1往復するデバイスへのアクセス時間と比較して短くできる。したがって、1つのOSに複数の同じデバイスを割り当てた場合に、デバイスに対する入出力命令のレイテンシ及びターンアラウンドタイムを短縮できる。

【0191】

また、デバイスセルに複数の同じデバイスを備え、デバイス管理情報で指定された処理を複数のデバイスのうちの任意のデバイスに実行させることで、例えば、任意のデバイスを利用している最中に障害が発生した場合、他のデバイスを利用して処理を続行できる。これにより、障害時のフェイルオーバーを実現できる。また、処理を複数のデバイスに分散させることもできる。

【0192】

さらに、CPUセルに、CPUから発行された複数の命令をまとめたコマンドを生成するコマンド送出回路を備え、デバイスセルに、コマンドを分解して複数の命令を抽出するコマンド解析部を備え、抽出された複数の命令により更新されるデバイス管理情報で指定された処理をデバイスに実行させることで、デバイスに対する入出力命令等によるネットワークのトラフィックが低減するため、デバイスに対するアクセス性能が向上する。

【0193】

また、CPUセルで、自機が所属する前記グループを特定するためのシステム識別子を保持し、デバイスセルで、システム識別子に対応するCPUセルのリストから構成されるシステム構成情報を保持し、デバイスセルは、コマンドと共にCPUセルから送信されたシステム識別子と保持したシステム構成情報から該CPUセルが所属するグループの中から任意の一つのCPUセルを選択し、該選択したCPUセルに対して該コマンドに対する処理結果を含む応答メッセージを送出し、CPUセルは、デバイスセルから応答メッセージを受信すると、該応答メッセージにしたがってデバイスセルの処理結果を取得することで、デバイスセル

で処理を行っている最中にシステム構成が変更されても正しく応答を返すことができる。また、CPUセルとデバイスセル間を接続するネットワークやCPUセルの負荷を分散することができる。

【図面の簡単な説明】

【図 1】

本発明のマルチプロセッサシステムの第 1 の実施の形態の構成を示すブロック図である。

【図 2】

図 1 に示したネットワーク上で伝送されるパケットの構成を示す模式図である。

【図 3】

図 1 に示したデバイスのポート構成を示すブロック図である。

【図 4】

図 1 に示したデバイス制御ブリッジが備えるデバイス管理情報の構成を示すブロック図である。

【図 5】

本発明の第 2 の実施の形態のマルチプロセッサシステムが有するデバイス管理情報の構成を示すブロック図である。

【図 6】

本発明のマルチプロセッサシステムの第 3 の実施の形態の処理手順を示すフローチャートである。

【図 7】

第 3 の実施の形態のマルチプロセッサシステムの各システムが備えるデバイスドライバが保持する情報を示すテーブル図である。

【図 8】

第 4 の実施の形態のマルチプロセッサシステムが有するデバイスセルの構成を示すブロック図である。

【図 9】

第 4 の実施の形態のマルチプロセッサシステムが有するデバイスセルの変形例

の構成を示すブロック図である。

【図 10】

図 9 に示したデバイス管理部が保持する情報を示すテーブル図である。

【図 11】

本発明のマルチプロセッサシステムの第 4 の実施の形態の処理手順を示すフローチャートである。

【図 12】

本発明のマルチプロセッサシステムの第 5 の実施の形態の構成を示すブロック図である。

【図 13】

図 12 に示したデバイス制御ブリッジが備えるデバイス管理情報の構成を示すブロック図である。

【図 14】

マルチプロセッサシステムの第 1 従来例の構成を示すブロック図である。

【図 15】

マルチプロセッサシステムの第 3 従来例の構成を示すブロック図である。

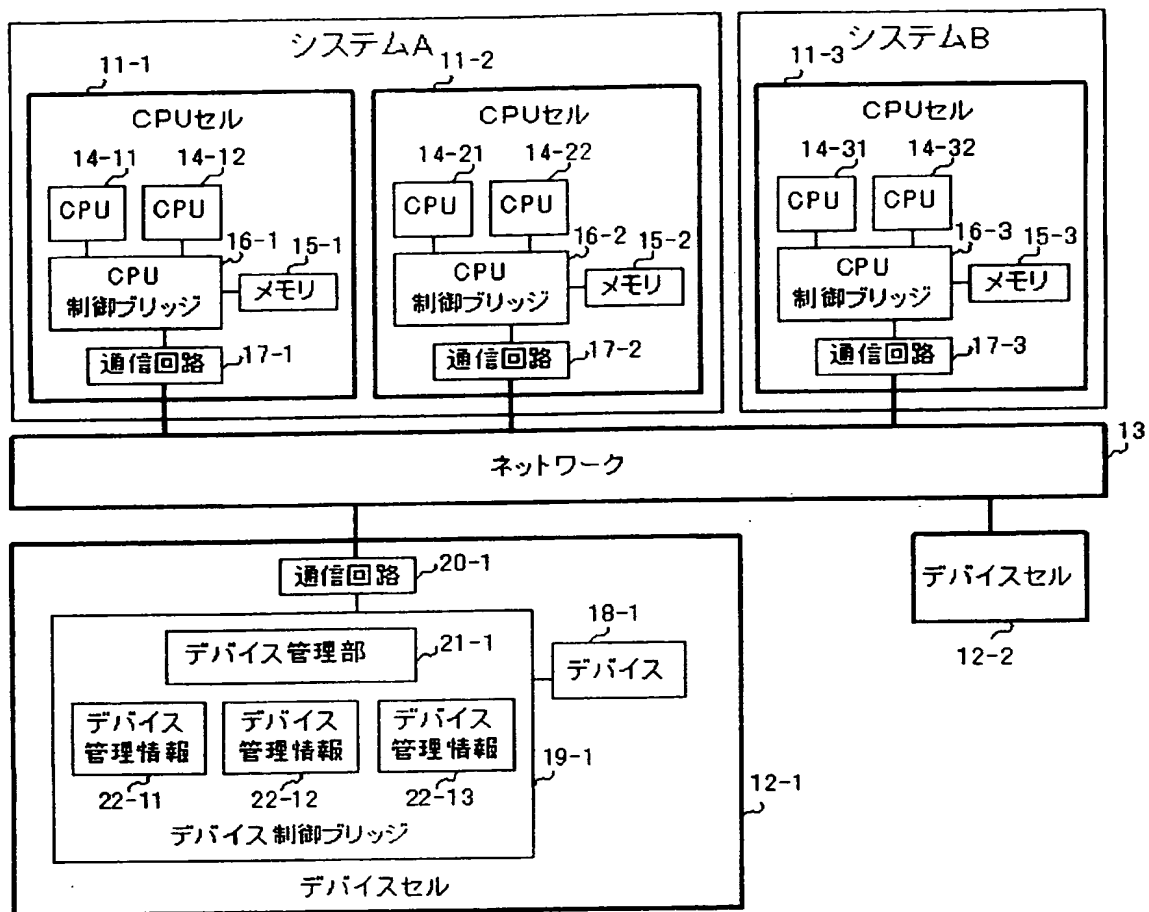
【符号の説明】

- 11 CPUセル
- 12 デバイスセル
- 13 ネットワーク
- 14 CPU
- 15 メモリ
- 16 CPU制御ブリッジ
- 17 通信回路
- 18 デバイス
- 19 デバイス制御ブリッジ
- 20 通信回路
- 21 デバイス管理部
- 22 デバイス管理情報

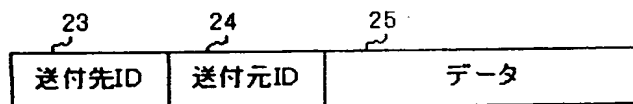
- 2 3 送付先 I D 領域
- 2 4 送付元 I D 領域
- 2 5 データ領域
- 2 6 モード設定ポート
- 2 7 処理設定ポート
- 2 8 結果読出ポート
- 2 9 モードレジスタ
- 3 0 結果レジスタ
- 3 1 C P U セルレジスタ
- 3 2 入出力ポートベースレジスタ
- 3 3 入出力ポート長レジスタ
- 3 4 C P U セル I D
- 3 5 入出力ポートベース
- 3 6、3 7 ロックビット
- 3 8 コマンド生成回路
- 3 9 コマンド解析部

【書類名】 図面

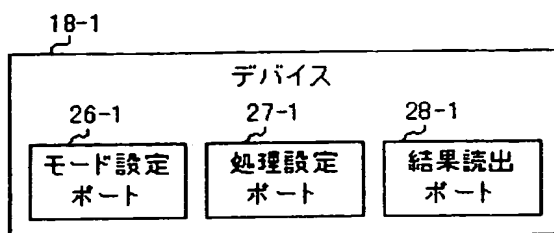
【図 1】



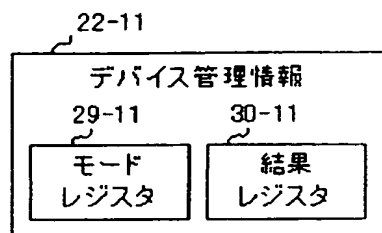
【図 2】



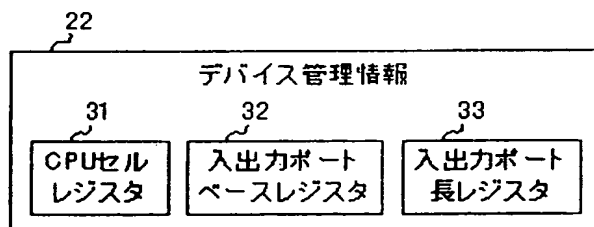
【図 3】



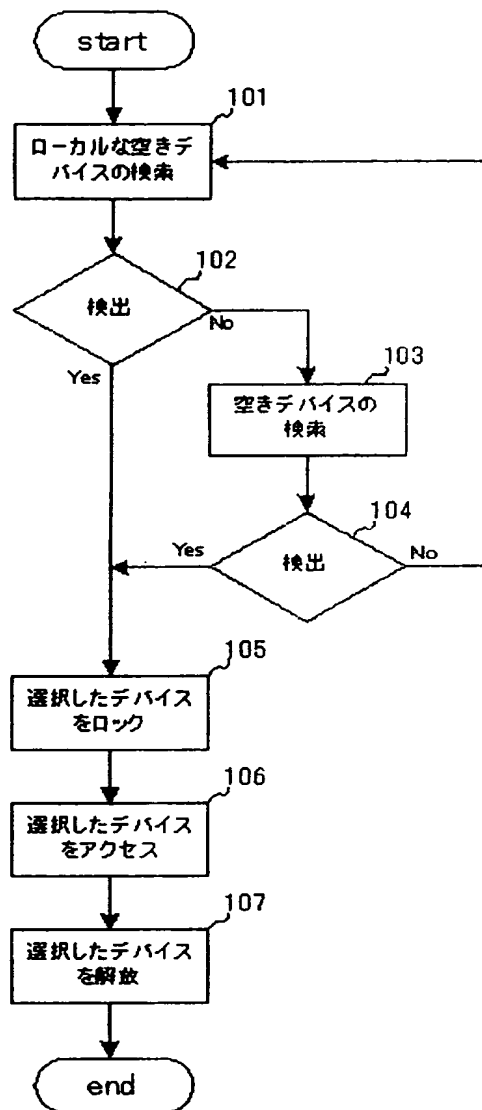
【図 4】



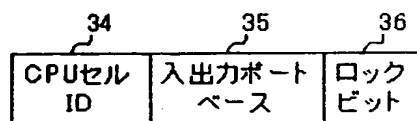
【図 5】



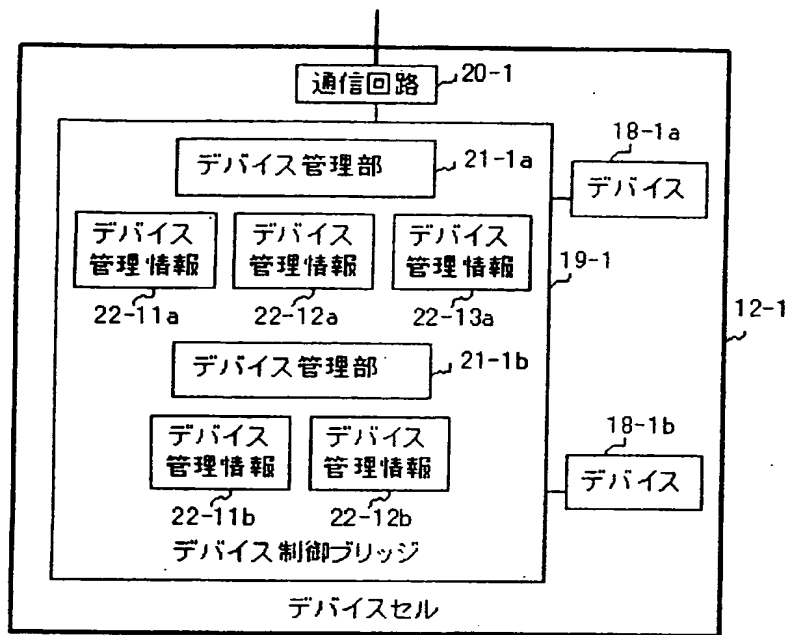
【図 6】



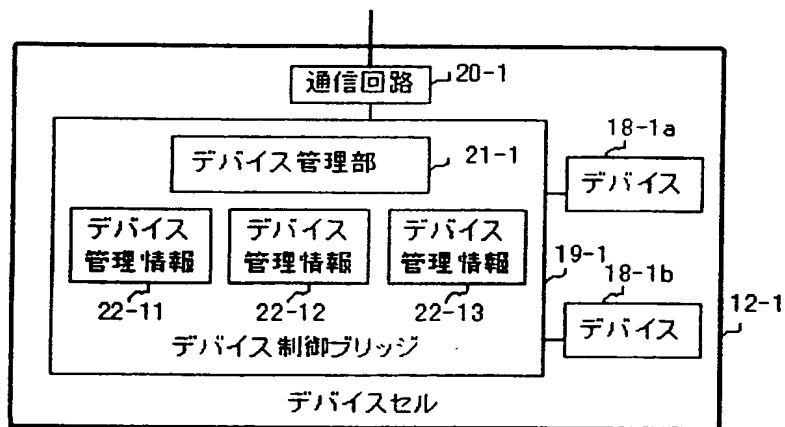
【図 7】



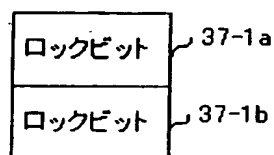
【図 8】



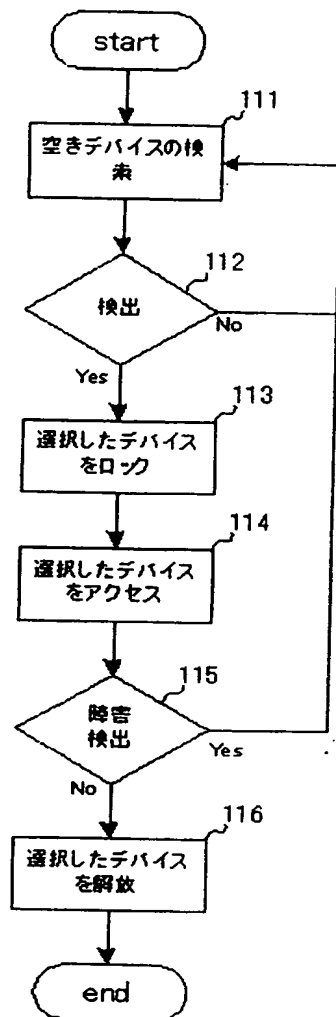
【図 9】



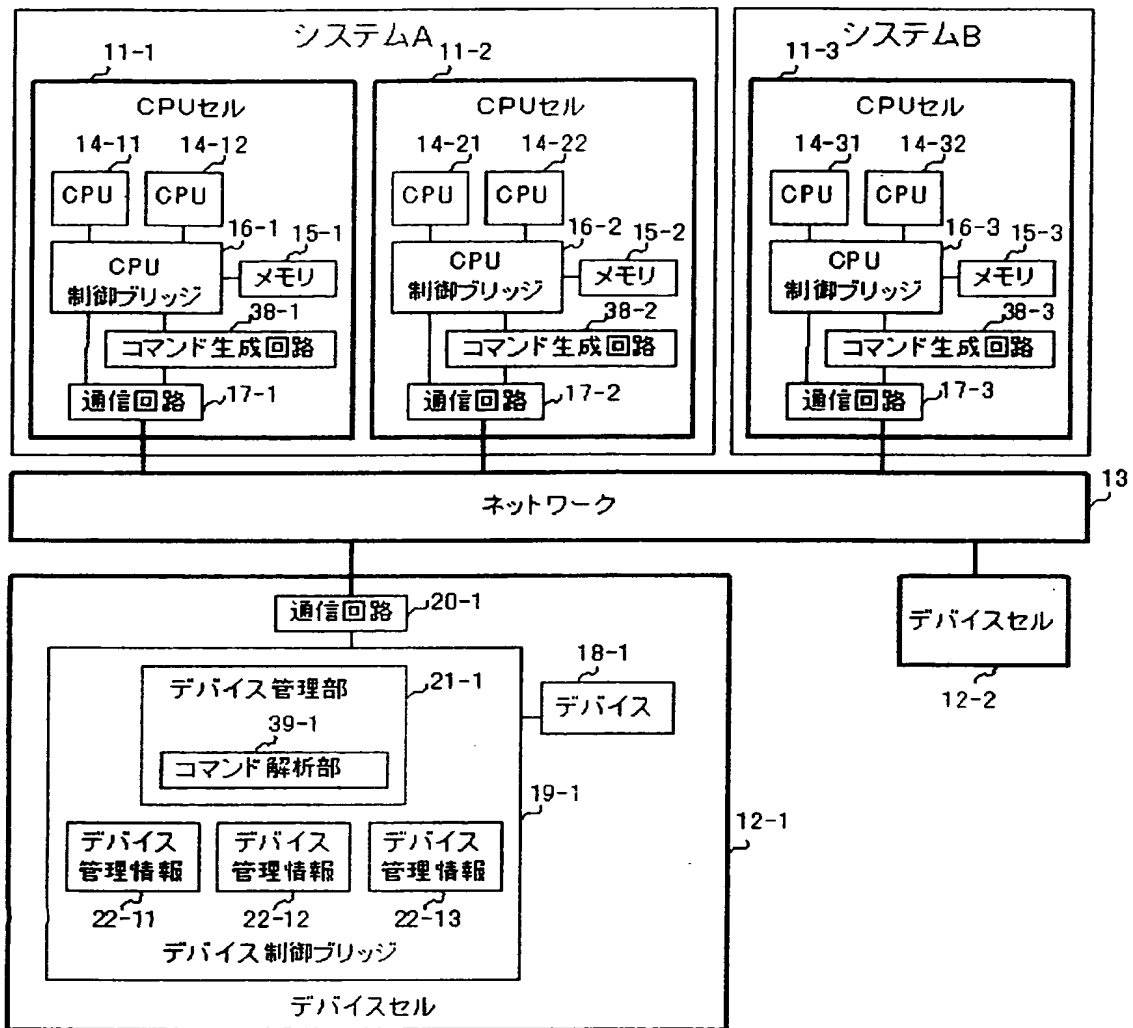
【図 10】



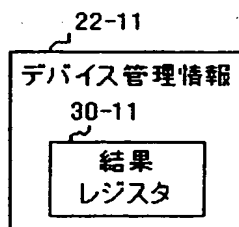
【図 11】



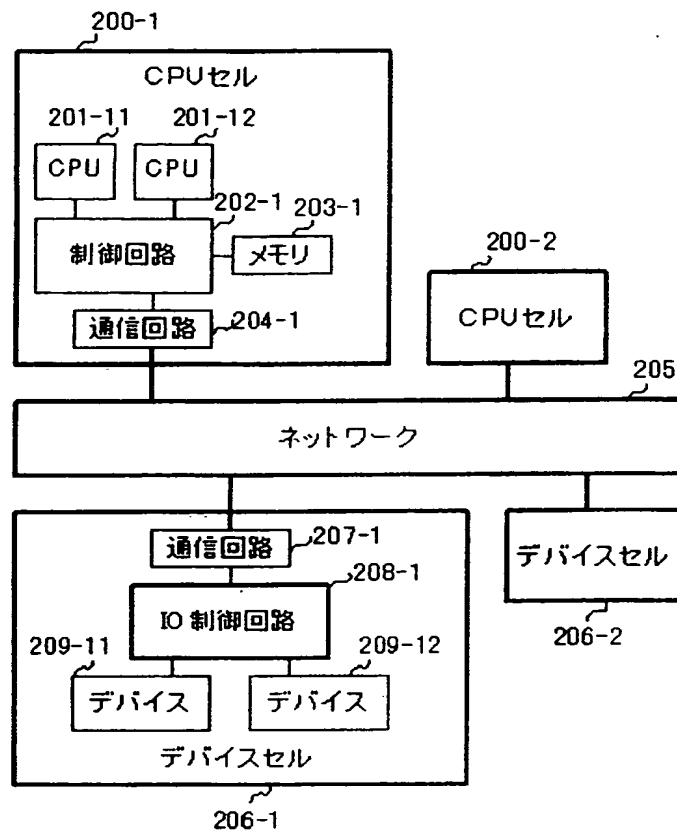
【図 1 2】



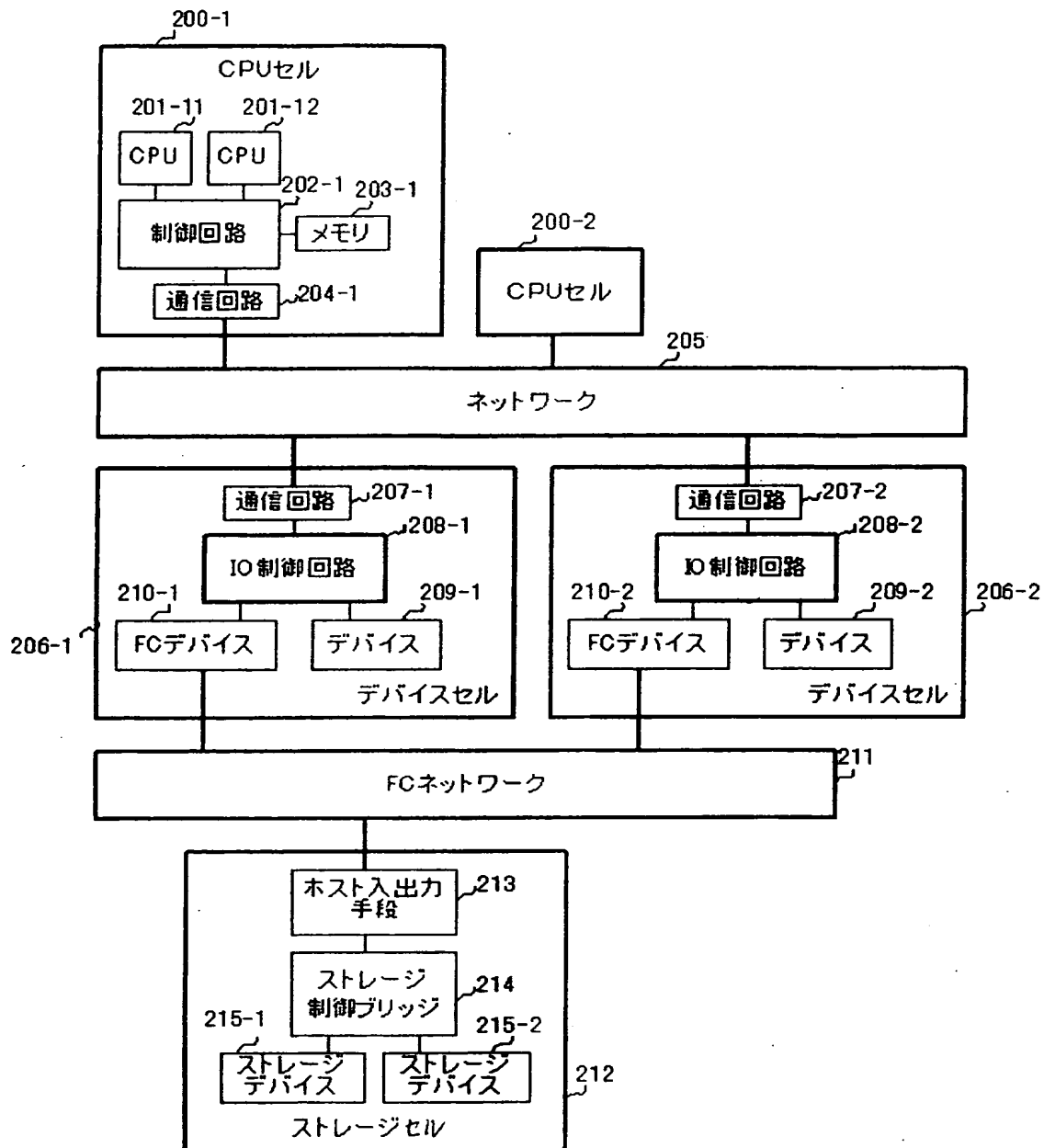
【図 1 3】



【図 14】



【図 15】



【書類名】 要約書

【要約】

【課題】 複数のOSで動作する環境下であってもデバイスを共有化することが可能な簡易な構成のマルチプロセッサシステムを提供する。

【解決手段】 少なくとも1つのCPUを備えた複数のCPUセルを有し、CPUセルが複数のグループに分割され、該グループ毎に異なるオペレーティングシステムで動作するマルチプロセッサシステムでデバイスを共用するためのデバイス共有方法であって、CPUセルとネットワークを介して接続されたデバイスを含むデバイスセルに、デバイスで実行可能な複数種類の処理を管理するためのデバイス管理情報を備えておき、デバイスセルでCPUセルからコマンドを受信すると、該コマンドの発行元に対応するデバイス管理情報を検索し、コマンドにより更新されるデバイス管理情報で指定された処理をデバイスに実行させる。

【選択図】 図1

特願 2003-038201

出 願 人 履 歷 情 報

識別番号

[000004237]

1. 変更年月日

1990年 8月29日

[変更理由]

新規登録

住 所

東京都港区芝五丁目7番1号

氏 名

日本電気株式会社